

The Assignment of Grammatical Gender in German: Testing Optimal Gender Assignment Theory

Emma Charlotte Corteen

Trinity Hall

September 2018

This dissertation is submitted for the degree of Doctor of Philosophy

The Assignment of Grammatical Gender in German: Testing Optimal Gender Assignment Theory

Emma Charlotte Corteen

Abstract

The assignment of grammatical gender in German is a notoriously problematic phenomenon due to the apparent opacity of the gender assignment system (e.g. Comrie 1999: 461). Various models of German gender assignment have been proposed (e.g. Spitz 1965, Köpcke 1982, Corbett 1991, Wegener 1995), but none of these is able to account for all of the German data.

This thesis investigates a relatively under-explored, recent approach to German gender assignment in the form of Optimal Gender Assignment Theory (OGAT), proposed by Rice (2006). Using the framework of Optimality Theory, OGAT claims that the form and meaning of a noun are of equal importance with respect to its gender. This is formally represented by the crucial equal ranking of all gender assignment constraints in a block of GENDER FEATURES, which is in turn ranked above a default markedness hierarchy *NEUTER » *FEMININE » *MASCULINE, which is based on category size.

A key weakness of OGAT is that it does not specify what constitutes a valid GENDER FEATURES constraint. This means that, in theory, any constraint can be proposed ad hoc to ensure that an OGAT analysis yields the correct result. In order to prevent any constraints based on ‘postfactum rationalisations’ (Comrie 1999: 461) from being included in the investigation, the GENDER FEATURES constraints which have been proposed in the literature for German are assessed according to six criteria suggested by Enger (2009), which seek to determine whether there is independent evidence for a GENDER FEATURES constraint.

Using an independently-verified constraint set, OGAT is then tested on a sample of 592 nouns systematically selected from the Duden *Rechtschreibung*. The results indicate that OGAT is relatively successful in its predictions when compared to other approaches but that it cannot account fully for the sample data. Accordingly, a revised version of the theory is proposed (OGAT II), which involves the ranking of certain GENDER FEATURES constraints. It is found that OGAT II is able to account for the genders of around 95% of nouns in the sample. A number of specific aspects of OGAT II are then tested by means of an experiment in which native German speakers are required to assign genders to 26 pseudo-nouns. The results suggest that OGAT II comes the closest of the systems discussed in the literature to modelling how native speakers assign gender in German.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. This dissertation does not exceed the prescribed word limit for the Degree Committee of the Faculty of Modern and Medieval Languages.

Acknowledgements

Firstly, I would like to thank the Arts and Humanities Research Council for funding this project. I would also like to thank the Cambridge Trust and the Pigott Scholars Programme for their offers of funding, and the Vice Chancellor for my Honorary Vice Chancellor's Award.

My greatest thanks go to my supervisor, Dr Sheila Watts, for her continual support and guidance throughout this project. I am extremely grateful not only for her invaluable academic insight, but also for her never-ending patience, positivity and encouragement.

I am also grateful to all of the other members of the Department of German and Dutch and the Department of Theoretical and Applied Linguistics at Cambridge University who provided me with advice and support throughout my time at Cambridge. Particular thanks go to Dr Bert Vaux, Dr Napoleon Katsos, Dr David Willis, Dr Jim Blevins, Professor Sarah Colvin and Dr Simon Pickl.

I would also like to thank the Cambridge Statistics Clinic at the Centre for Mathematical Sciences for their assistance with my statistical queries.

My thanks also go to the many native German speakers who I consulted throughout the project, and particularly the 100 who took part in my experiment.

Additionally, I would like to thank members of staff at the Schools of Modern Languages and English Literature, Language and Linguistics at Newcastle University, where my studies of Linguistics began. In particular, I would like to thank Dr Carol Fehringer and Andrea Wilczynski for nurturing my interest in German linguistics and encouraging me to apply for postgraduate study.

I would also like to thank my friends and family for their support throughout the PhD writing process. Thank you to those of you who at least sometimes pretended to be interested in German grammar.

Last but by no means least, my thanks go to Eve, without whom I would not have been able to complete this project.

Contents

Chapter 1: Introduction	1
1.1 Grammatical gender in German	1
1.2 Gender assignment in German	2
1.3 Gender assignment principles	7
1.3.1 Morphological principles	7
1.3.1.1 Derivation via affixation.....	8
1.3.1.2 Implicit derivation.....	9
1.3.1.3 Conversion	9
1.3.1.4 Compounds.....	10
1.3.1.5 Reductions	10
1.3.1.6 Inflection	11
1.3.1.7 Evidence for morphological principles.....	14
1.3.2 Phonological principles	15
1.3.2.1 Word ending	15
1.3.2.2 Other phonological factors	16
1.3.2.3 Evidence for phonological principles	17
1.3.3 Semantic principles	18
1.3.3.1 The Natural Gender Principle (NGP).....	18
1.3.3.2 Other semantic classes	19
1.3.3.3 Levels of abstraction	20
1.3.3.4 Evidence for semantic principles	22
1.4 Interaction of gender assignment principles	23
1.4.1 Semantics » form	24
1.4.2 Form » semantics.....	26
1.4.3 Morphology » semantics » phonology	27
1.4.4 Hierarchy of individual principles	28
1.4.5 Optimal Gender Assignment Theory (OGAT).....	29
1.4.5.1 Theory outline	29
1.4.5.2 Strengths of OGAT	34
1.4.5.3 Weaknesses and aspects of OGAT for further investigation	37
1.4.5.4 Previous testing of OGAT	45
1.4.5.4.1 OGAT and German	45
1.4.5.4.2 OGAT and other languages	49
1.5 Summary and plan for investigation	52
Chapter 2: Evaluating GENDER FEATURES constraints	55
2.1 Introduction	55
2.2 The noun sample and the constraints.....	56
2.3 Criterion 1: Coverage of candidate nouns	60
2.4 Criterion 2: Typological parallels.....	70
2.5 Criterion 3: Parallels outside of gender assignment.....	78
2.6 Criterion 4: Clarity of semantic features.....	85

2.7 Criterion 5: Accounting for exceptions	91
2.8 Criterion 6: Productivity	95
2.8.1 Neologisms.....	97
2.8.2 Historical gender changes.....	100
2.9 Refinement of constraint set	103
2.10 Summary of results	109
Chapter 3: Analysis – testing OGAT	113
3.1 Introduction and preliminary analysis	113
3.2 Constraint-related issues	118
3.2.1 Constraints with multiple versions	118
3.2.2 Constraints whose application is problematic	121
3.2.2.1 DEVERBAL STEM→*F,*N	121
3.2.2.2 SUPERORDINATES→*M,*F	129
3.2.2.3 DEVERBAL -T→*M,*N	135
3.2.2.4 -MA→*M,*F	137
3.3 Noun-related issues	138
3.3.1 Loanwords.....	138
3.3.1.1 Donor gender	140
3.3.1.2 Semantic equivalence	141
3.3.1.3 Homonymy/partial homonymy	147
3.3.1.4 Default gender	149
3.3.1.5 Summary	150
3.3.2 Nouns with multiple meanings or forms	150
3.3.2.1 Meaning variation.....	151
3.3.2.1.1 Multiple meanings vs. one meaning	151
3.3.2.1.2 Multiple meanings vs. one meaning in the sample	153
3.3.2.1.3 Polysemy vs. homonymy	153
3.3.2.1.4 Polysemy vs. homonymy in the sample	156
3.3.2.1.5 Approaches to polysemy	157
3.3.2.1.6 Polysemy in the sample	160
3.3.2.2 Form variation	162
3.3.3 Complex nouns and “strong rules”	163
3.3.3.1 Heads	164
3.3.3.1.1 Head identification	164
3.3.3.1.2 Head position.....	165
3.3.3.2 Multi-word units and OGAT.....	166
3.3.3.2.1 Compounds.....	166
3.3.3.2.2 Lexicalised phrases	167
3.3.3.2.3 Multi-word units and OGAT.....	168
3.3.3.3 Nouns derived via affixation and OGAT.....	169
3.3.4 Subset of 55 nouns	173
3.3.4.1 Placenames	173
3.3.4.2 Personal names	178

3.3.4.3 Plural terms	181
3.3.4.4 Regional terms	182
3.3.4.5 Archaic nouns	183
3.3.4.6 Homonyms	184
3.3.4.7 Adjectival nouns	185
3.3.4.8 Multiple-gender nouns	186
3.4 Remaining issues.....	191
3.4.1 New constraints	192
3.4.2 Other nouns in the sample	196
3.4.3 Testing alternative default hierarchies	198
3.5 Effectiveness of OGAT vs. OGAT II and exceptions	200
3.5.1 OGAT vs. OGAT II	200
3.5.2 Exceptions	201
3.5.3 OGAT II vs. other approaches	204
3.5.4 Economising the theory	205
3.6 Summary	207
Chapter 4: Testing the predictive power of OGAT II	209
4.1 Introduction	209
4.2 Experiment design	209
4.3 Method	213
4.4 Results	216
4.5 Discussion	219
4.6 Conclusions and implications of results.....	232
Chapter 5: Conclusions and directions for further work	241
Reference list	247
Appendices.....	273
Appendix A: The sample	273
Appendix B: The constraints	277
Appendix C: Sources for criterion 1	303
Appendix D: Sources for criterion 4.....	309
Appendix E: Summary of results from Chapter 2	313
Appendix F: Summary of results from Chapter 3	319
Appendix G: Experiment	339

Chapter 1: Introduction

'A person who has not studied German can form no idea of what a perplexing language it is [...]. Every noun has a gender, and there is no sense or system in the distribution; so the gender of each must be learned separately and by heart. There is no other way. To do this one has to have a memory like a memorandum-book. In German, a young lady has no sex, while a turnip has. Think what overwrought reverence that shows for the turnip, and what callous disrespect for the girl.'

– Mark Twain, *The Awful German Language* (1879)

1.1 Grammatical gender in German

Gender is a grammatical feature present in around half of the world's languages which separates nouns into classes (Corbett 2013, Audring 2016). In German, there are three grammatical genders: masculine, feminine and neuter. Almost all German nouns belong to one of these categories. For instance, *Tisch* 'table' is masculine, *Lampe* 'lamp' is feminine, and *Fenster* 'window' is neuter. In German, gender is said to be 'covert' because the gender of a noun is not marked on the noun (the "controller") itself (Corbett 1991: 62). In accordance with Hockett's (1958: 231) well-known definition of gender, 'genders are classes of nouns reflected in the behaviour of associated words', gender in German is visible only on agreeing elements (known as "targets") such as determiners, attributive adjectives, relative pronouns and personal pronouns. For example, the masculine gender of *Tisch* is visible on the determiner in *der Tisch* 'the.MASC table'.

The gender of a German noun can, however, only be seen on agreeing targets when the noun is in the singular. In the plural, gender is said to be 'neutralised', since all targets are marked identically (e.g. Bauch 1971: 412, Weinrich 2007: 326, Kürschner & Nübling 2011: 363). Accordingly, the small subset of nouns which appear only in the plural and have no singular variant, known as *pluralia tantum*, are said to have no gender, e.g. *Eltern* 'parents', *Ferien* 'holidays', *Leute* 'people' (e.g. Wegener 2000: 541, Nübling 2008: 289).

For all nouns with a singular form, however, i.e. most German nouns, gender is obligatory (e.g. MacKay 1999: 73, Hickey 2000: 639). The vast majority of German nouns belong unequivocally to just one gender. There are, however, two main categories of nouns for which this is not the case: "multiple-gender" nouns and "hybrid" nouns.

So-called "multiple-gender" nouns (MGNs) are nouns which can be used with more than one gender without any effect on the noun's meaning, such as *der/das Sims* 'window sill' (Corbett 1991: 181). MGNs are therefore not to be confused with homonyms such as *die Leiter* 'ladder' and *der Leiter* 'leader'. The gender of a MGN may vary according to factors such as region, e.g. *die Butter* (Standard German) vs. *der Butter* (Southern Germany and Austria), or register, e.g. *der Virus* (everyday speech)

vs. *das Virus* (scientific contexts) (e.g. Hickey 2000: 625, Duden Online-Wörterbuch 2018f). Many MGNs can also be found amongst relatively recent loanwords, e.g. *der/das Laptop* (e.g. Hoberg 2004: 79). MGNs constitute just 1.3% of nouns listed in the Duden dictionary (the standard reference work for German) (Duden Sprachwissen 2017).

“Hybrid” nouns, are nouns which ‘neither [...] take the agreements of one consistent agreement pattern nor belong to two or more genders’ (Corbett 1991: 183). That is to say, while the hybrid noun itself is said to belong to just one gender, its targets may display agreement with more than one gender. In German, hybrid nouns are usually those which refer to people but whose grammatical gender does not match the conceptual gender of the referent, e.g. *Mädchen* ‘girl’. In line with Corbett’s (1979, 1991, 2006) Agreement Hierarchy, all attributive targets used in conjunction with *Mädchen* agree with the noun’s grammatical gender (neuter), but anaphoric personal pronouns may agree with either the noun’s grammatical gender or the conceptual gender of the referent (feminine) (e.g. Batliner 1984: 850, Semplicini 2012: 141). It has been argued that the kind of agreement displayed by an anaphoric personal pronoun may be influenced by factors such as the distance between the controller noun and the personal pronoun (Köpcke, Panther & Zubin 2010) or, in the case of *Mädchen* specifically, the age of the referent in question (Braun & Haig 2010). Due to the presence of hybrid nouns in German, it is important to look at attributive targets only in order to reliably ascertain the grammatical gender of a noun.

1.2 Gender assignment in German

The question of what determines the grammatical gender of a noun is highly intriguing, and is one which is central to the present study. How is it that native speakers know which gender a particular noun should be? Why, for instance, is *Schuh* ‘shoe’ masculine but *Mütze* ‘cap’ feminine? In some languages, the reasons for nouns being allotted to different genders is relatively transparent, e.g. Tamil (Corbett 1991: 9) and Russian (Rodina & Westergaard 2013: 95). German, however, is notorious for its seemingly opaque gender system (e.g. Salmons 1993: 187, Comrie 1999: 461, Hickey 2000: 628, Schmid 2002: 100), which has led various linguists to claim that German nouns are assigned genders completely arbitrarily.

The notion of the arbitrariness of the German gender system was widespread in the early 20th century. Bloomfield (1933: 271, 280), for instance, states that ‘there seems to be no practical criterion by which the gender of a noun in German [...] could be determined’. This stance has also been maintained by some linguists in more recent times. Maratsos (1979: 235), for instance, states that ‘the classification is arbitrary, [...] no underlying rationale can be guessed at’, and, even more recently, Pfau (2009: 109)

declares that ‘in German, [...] the gender of only a very small percentage of nouns can be predicted [...], that is, the German gender system is largely arbitrary’.

According to this view, due to the perceived lack of regularity in gender assignment, speakers have no option but to simply memorise the gender of each noun. That is to say, gender must be stored for each noun in the lexicon and retrieved each time a noun is used (Corbett 1991: 7, van Berkum 1996: 33, Conzett 2006: 224).

However, over the past few decades, results from a number of in-depth analyses based on corpus data have revealed various patterns of correlation between the gender of a German noun and its semantic, morphological and phonological properties (e.g. Arndt 1970, Altmann & Raettig 1973, Zubin & Köpcke 1981, 1986, Köpcke 1982, Köpcke & Zubin 1983, 1984). These patterns will be outlined in Section 1.3.

Studies have proposed that using the regularities observed, the gender of a large proportion of German nouns can be systematically predicted. For instance, Zubin & Köpcke (1981, 1986), Köpcke (1982), and Köpcke & Zubin (1983, 1984) estimate that the gender of around 90% of the 1466 monosyllables in the Duden can be accounted for using a combination of semantic, phonological and morphological principles.¹ Such findings have led many linguists to conclude that gender assignment in German is not arbitrary but rather a largely rule-governed process (e.g. Mills 1986, Wegener 1995, Eisenberg 1999, Schmid 2002).

Furthermore, similar estimates have been made for other languages whose gender classification has also previously been dismissed as arbitrary. For Dutch, for instance, Frieke (1988) demonstrates the ability of a set of gender assignment rules to account for the genders of around 80% of a corpus of 763 monosyllables. For Norwegian, Trosterud (2001) proposes a set of assignment rules which he claims cover around 94% of nouns, and for French, Tucker, Lambert & Rigault (1977) establish the ability of a set of phonological gender assignment rules to account for around 85% of a corpus of 31,619 nouns.²

On the basis of such data, Corbett (1994: 1350) argues that gender assignment is ‘essentially systematic’ in all languages. He reasons that, due to the high degree of regularity in gender assignment, native speakers do not have to rely on the memorisation of gender (i.e. lexical storage), but can instead compute the gender of the vast majority of nouns based on their semantic, morphological and phonological information, which is in any case available in the lexicon (Corbett 1991: 66ff.). That is to say, speakers possess a ‘gender assignment system’, specifically in the lexicon,

¹ There are, however, a small number of issues with their proposed model, which will be discussed in 1.4.

² See Corbett (1991) for further data from studies of numerous other languages.

which enables them to derive the genders of most nouns based on their formal and semantic features (Corbett 1991: 7, 66). Speakers therefore must only store the genders of the small proportion of nouns whose gender is synchronically unpredictable (e.g. Mills 1986: 41, Hickey 2000: 644).

Furthermore, Corbett (1991: 7), alongside other linguists, argues in favour of the computation of gender on the basis that the alternative, namely storage, is problematic. He states that having to simply remember the gender of every single noun in the lexicon without taking advantage of any regularities would entail 'a considerable feat of memory'. Zubin & Köpcke (1981: 447) and Corbett (1991: 7) reason that if the gender of every noun had to be memorised, then it would be expected that native speakers would make frequent gender errors (as with non-natives), which they do not. Moreover, it is argued that the memorisation of the genders of tens of thousands of nouns would present an 'insurmountable task to the language learner' (Zubin & Köpcke 1981: 447) and would place 'extreme stress' on a learner's memory (Wegener 2000: 514).

However, van Berkum (1996: 31f.) states that this argument alone is not sufficient to dismiss the possibility of gender storage. He argues that, if native speakers are able to learn the 'essentially arbitrary' form of each item in the lexicon, then it is not impossible that they would be able to do the same for gender. Van Berkum (1996: 32f.) claims that native speakers may well both have a system for deriving gender online and store gender too. He does not doubt that native speakers of Dutch are able to derive the neuter gender of *meisje* 'girl' from the fact that it is a diminutive, but maintains that this does not eliminate the possibility of storing the gender of this noun as well. Although systematicity in gender assignment would undeniably aid acquisition, van Berkum (1996: 32) argues that there is no a priori reason to assume that gender could not be learned without it.

As noted by e.g. Conzett (2006: 230), the perceived validity of this argument is highly dependent on one's view of redundancy in the language faculty. To have a system of regularities but not always exploit them or not exploit them at all would be highly redundant and would violate the linguistic principles of economy and parsimony, as well as the more general principle of Occam's razor. As Salmons (1994: 187) states, to say that German gender assignment is not rule based is to violate the claim that 'language constitutes rule-governed behaviour'. However, the premise of this argument is not universally held, and, given the scale of the issue, its discussion is not within the scope of this project.

A further argument cited in favour of the computation of gender is the fact that nouns which have been borrowed into German from other languages necessarily acquire a gender (e.g. Hickey 2000: 639), which demonstrates that there must be a mechanism for assigning and not just remembering gender (Corbett 1991: 7). Moreover, as observed by e.g. Hock & Joseph (2009: 254), native speakers

generally tend to agree on the gender that should be assigned to a loan, suggesting that this mechanism is part of the linguistic competence of all native speakers. This is corroborated by the results from a number of studies on loanwords in German, e.g. Arndt (1970), Carstensen (1980), Gregor (1983), Callies, Onsyko & Ogiermann (2012) etc.

Similarly, when presented with pseudo-nouns in an experimental setting, i.e. a scenario in which gender retrieval is simply not a possibility, speakers are able to assign genders to the pseudo-nouns (e.g. Köpcke & Zubin 1983: 167, Corbett 1991: 7). As with loans, speakers generally do this with a high degree of consistency, as is illustrated by e.g. Lang (1976), Carstensen (1980), Köpcke & Zubin (1983), Mills (1986), Wegener (1995, 2000), Schwichtenberg & Schiller (2004) and Hohlfeld (2006) for German; Mastropavlou & Tsimpli (2011) for Greek; Zekhnini & Hulstijn (1995) for Dutch; Tucker et al. (1968, 1977) and Karmiloff-Smith (1979) for French; and Sadek, Kiraithe & Villareal (1975) for Spanish. This suggests that native speakers have the ability to somehow “work out” the gender of a noun based on its features (Corbett 1991: 7).

This is generally perceived to be convincing supporting evidence for the presence of a mechanism in the language faculty that is able to assign gender (e.g. Conzett 2006: 236). However, van Berkum (1996: 31) highlights that, while it is possible that there is indeed a mechanism in the language faculty which is used to assign gender to “new” words (i.e. loans, pseudo-nouns, previously unheard native nouns), it is also possible that this mechanism is not always exploited or not exploited at all for assigning gender to known words, i.e. that gender is derived for new nouns but stored for existing nouns.

Again, the validity of this counter-proposal is similarly dependent on one’s view of redundancy in the language faculty. While it is possible that a native speaker possesses a mechanism that can assign gender to the vast majority of nouns but does not always exploit it unless the noun is new, it would be a highly redundant model of gender assignment. This is especially true because presumably the mechanism for assigning gender to new nouns would consist of generalisations which have been made based on existing nouns in the lexicon. Nevertheless, even if the system were redundant and the mechanism were not exploited for all nouns, it is still agreed that a mechanism for gender assignment exists, and it is this mechanism which is of interest.

In spite of van Berkum’s (1996: 33) criticisms of the arguments in favour of gender computation, he acknowledges that, in light of the high degree of regularity observed by e.g. Frieke (1988), Zubin & Köpcke (1981) and Corbett (1991), the computation of gender is still very much a possibility and one which merits further exploration. He states that, ultimately, the only conclusive evidence can come

from empirical testing in order to determine whether native speakers pick up on and make use of the regularities observed, a point also raised by Corbett (1991: 70) and Conzett (2006: 229).

Psycholinguistic experiments designed to test this very question, i.e. whether native speakers take advantage of gender assignment regularities or not, have produced very mixed results, most likely due to the wide variety of experimental methods used.

On the one hand, for example, results from experiments by a number of linguists on speakers who are anomic or in a “tip-of-the-tongue” state, show that the participants were able to correctly assign gender to nouns without being able to retrieve their word form, suggesting that gender can be assigned without knowledge of a noun’s phonological form. This was found for e.g. Italian by Badecker et al. (1995), Miozzo & Caramazza (1997), Caramazza & Miozzo (1997) and Vigliocco et al. (1997); for French by Henaff Gonon et al. (1989); and for Spanish by Avila et al. (2001). These linguists therefore conclude that formal features cannot influence gender assignment, arguing in favour of the gender storage model. However, Conzett (2006: 231ff.) highlights some potential methodological issues with a number of these studies, and, moreover, the results of these experiments are not incompatible with the idea that semantic features may have an effect on gender assignment.

On the other hand, results from a number of other psycholinguistic studies contradict the findings of these experiments. For instance, in their study on Italian native speakers, Caffarra et al. (2015) find ERP (event-related brain potential) data which demonstrates that the brain detects formal cues to gender during sentence processing in line with the gender computation model, i.e. that noun ending does have an effect on gender assignment in Italian. Additionally, results relating to error rates and reaction times from a number of experiments involving gender decision tasks provide evidence which supports the idea that native speakers make use of a noun’s formal features (specifically its ending) when assigning gender. This has been shown for French by e.g. Desrochers & Paivio (1990) and Taft & Meunier (1998); for Spanish by Hernandez et al. (2004) and Afonso et al. (2014); and for Italian by Padovani & Cacciari (2003) and De Martino et al. (2010).

Furthermore, van Berkum (1996: 93-187), in his study of Dutch gender, tests a specific model of gender storage, namely that proposed by Jescheniak & Levelt (1994), who state that gender is stored for all nouns and the retrieval of a noun’s gender from the lexicon is made faster by the gender of that noun having been retrieved recently before,³ and finds that his results contradict this hypothesis. While his results do not necessarily disconfirm that gender is stored, they do rule out this particular model of storage (van Berkum 1996: 187).

³ See Jeschniak & Levelt (1994) and van Berkum (1996: 95-100) for further details.

A third option has also emerged from the psycholinguistic literature, namely that gender can be both stored and computed. Bates et al. (1995), for example, found that Italian native speakers were influenced by word-ending in a gender selection task but not in a noun repetition task. Consequently, they propose that the gender of a noun can be retrieved independently of its form but then can also be computed on the basis of its formal features as a “checking” mechanism when the task requires. In their study of Hebrew, Gollan & Frost (2001) also find that computation as a method of “checking” occurs in specific tasks, such as when participants are presented with ungrammatical data. They state that the degree to which gender is computed is likely to be language specific, as there will be ‘differences in the availability and strength of [...] mechanisms for deriving gender’ (Gollan & Frost 2001: 648).

It therefore seems that, in spite of the conflicting evidence, the computation of gender remains a plausible possibility. However, the degree to which it occurs in a language and the circumstances under which it occurs are matters for further psycholinguistic investigation.

Given that there is at least some psycholinguistic evidence to suggest that native speakers do exploit a mechanism for computing gender, it seems worth exploring what this mechanism consists of. The next section will focus on the case of German specifically, and the various morphological, phonological and semantic gender assignment principles that have been proposed in the literature. For each type of principle, evidence for their employment by native German speakers will be discussed.

1.3 Gender assignment principles

Among those who maintain that gender in German is largely computed rather than stored, it is widely agreed that the mechanism for assigning gender uses regularities relating to a noun’s form and meaning (e.g. Zubin & Köpcke 1981, Köpcke 1982, Mills 1986, Corbett 1991, Wegener 1995, Köpcke & Zubin 1996, Hickey 2000, Hoberg 2004, Weinrich 2007). As a result of numerous studies on German gender, various morphological, phonological and semantic gender assignment principles have been formulated. In this section, a number of these principles will be discussed and evidence for their usage by native speakers will be presented.

1.3.1 Morphological principles

Various gender assignment principles have been proposed for German which are based on the morphology of a noun. These include principles relating to particular derivational affixes, to compounds, and even to the inflectional class of a noun. Several of these principles will be outlined below.

1.3.1.1 Derivation via affixation

The connection between a noun's derivational morphological features and its gender is well established for German (e.g. Arndt 1970, Mills 1986: 30f., Köpcke & Zubin 1996: 28, Hickey 2000: 630, Hoberg 2004: 86-89, Weinrich 2007: 326f.). Perhaps the most widely cited gender assignment principles are those involving derivational affixes, which typically categorically determine a noun's gender. For example, nouns with the suffixes *-er* (e.g. *Lehrer*) or *-ling* (e.g. *Bratling*) are masculine, nouns with the suffixes *-in* (e.g. *Lehrerin*), *-schaft* (e.g. *Freundschaft*) or *-ung* (e.g. *Meinung*) are feminine, and nouns with the suffixes *-chen* (e.g. *Bäumchen*) or *-lein* (e.g. *Büchlein*) are neuter (e.g. Altmann & Raettig 1973: 302, Flämig 1991: 453, Köpcke 1982: 71, Mills 1986: 30, Steinmetz & Rice 1989: 165, Köpcke & Zubin 1996: 476, Zifonun et al. 1997: 33, Hickey 2000: 630, 635, Hoberg 2004: 86f., Menzel 2004: 66, Thomoglou 2004: 41, Weinrich 2007: 326f., Duden Grammatik 2009: 164ff., Engel 2009: 280f., Eisenberg 2013: 133).

There are also, however, a small number of suffixes which are associated with two genders. For instance, nouns in *-nis* can be either neuter or feminine, e.g. *das Ereignis*, *die Erlaubnis*, and nouns in *-tum* can be neuter or masculine, e.g. *das Eigentum*, *der Reichtum* (e.g. Steinmetz 1986: 200, 2006: 1424; Flämig 1991: 453; Zifonun et al. 1997: 33; Nelson 1998: 215; Hickey 2000: 629; Hoberg 2004: 87ff.; Menzel 2004: 66; Thomoglou 2004: 41, Duden Grammatik 2009: 166, Engel 2009: 279, 282; Eisenberg 2013: 133). Nevertheless, for both of these suffixes, their association with neuter is stronger. Estimates of the proportion of neuter nouns in *-nis* range from around 60% (Hoberg 2004: 89) to over 75% (Leiss 2005: 23), and for nouns in *-tum*, all but two (*Reichtum* and *Irrtum*) of the 118 nouns listed in the Duden are neuter. Moreover, it is claimed that both *-nis* and *-tum* are now only productive in the neuter (e.g. Hickey 2000: 629, Eisenberg 2013: 134).

Additionally, a gender assignment principle has been proposed which involves the derivational prefix *Ge-*, stating that nouns with this feature are usually neuter, e.g. *Gefühl*, *Gestein* (e.g. Mills 1986: 30, Hickey 2000: 631, Menzel 2004: 66, Rice 2006: 1396, Steinmetz 2006: 1424). Unlike the majority of principles relating to suffixes, however, this principle is not categorical, but rather based on a tendency.

Also cited among the morphological gender assignment principles relating to affixes are those associating *Fremdsuffixe* (loan or "foreign" suffixes) with a particular gender. As with the principles relating to native suffixes, these are mostly (but not all) categorical. For example: nouns with the "foreign" suffixes *-ismus* (e.g. *Alkoholismus*) or *-ant* (e.g. *Protestant*) are masculine, nouns with the "foreign" suffixes *-anz* (e.g. *Arroganz*), *-enz* (e.g. *Existenz*), or *-(i)tät* (e.g. *Realität*) are feminine, and

nouns with the “foreign” suffixes *-ment* (e.g. *Fundament*) or *-ium* (e.g. *Studium*) are neuter (e.g. Altmann & Raettig 1973: 302, Mills 1986: 30, Steinmetz 1986: 197, Flämig 1991: 453, Zifonun et al. 1997: 33, Kratochvilová 2000: 71, Hoberg 2004: 86f., Menzel 2004: 66, Thomoglou 2004: 41, Chan 2005: 52, Weinrich 2007: 326f., Duden Grammatik 2009: 164ff., Engel 2009: 279ff.).

The gender of nouns derived via affixation will be considered further in Section 3.3.3.3.

1.3.1.2 Implicit derivation

In contrast to nouns derived by means of affixation, also termed “explicit” derivation (e.g. Fleischer & Barz 1992: 46, Donalies 2002: 97f.), there are also nouns which are claimed to have been formed via derivation without the use of affixes, a process termed “implicit” derivation (e.g. Fleischer & Barz 1992: 51, Russ 1994: 195). Nouns which are most commonly categorised as products of implicit derivation are nominalised verb stems with an ablauted stem vowel, e.g. *Gang*, *Flug*, *Zwang* (e.g. Fleischer & Barz 1992: 41, 2012: 89; Eisenberg 1999: 281; Donalies 2002: 98, 135; Lohde 2006: 50f.). Such nouns are relevant to gender assignment, since it has been observed that they have a strong association with masculine (Flämig 1991: 453, Hickey 2000: 643f., Hoberg 2004: 90, Chan 2005: 55).

Also sometimes categorised alongside these nouns are nominalised verb stems which have not undergone a stem vowel change, e.g. *Bau*, *Lauf*, *Versuch* (e.g. Russ 1994: 201). In terms of gender assignment, such a grouping is fitting, since it has been observed that these nouns are also typically masculine (e.g. Flämig 1991: 453, Hickey 2000: 643f., Hoberg 2004: 90, Chan 2005: 55). However, nominalised verb stems without ablaut are sometimes categorised as products not of implicit derivation but of conversion (e.g. Eisenberg 1999: 281, Hoberg 2004: 89), a process which will be discussed in the following section. Note that nouns of this type will be discussed in further detail in 3.2.2.1.

1.3.1.3 Conversion

Like implicit derivation, conversion is a word-formation process which does not involve (overt) affixation. It is usually seen either as being completely distinct from derivation, thought to consist of a simple change of syntactic category of a lexical item to form a new lexeme (e.g. Bauer 1983: 32, Plag 1999), or it is classified as a kind of derivation, often labelled “zero derivation”, and is thought to involve the affixation of a phonologically null, class-changing morpheme, or “zero affix”, to the base (e.g. Marchand 1969, Kiparsky 1982).

However the process is classified, it is widely claimed that words of certain categories which have been nominalised by means of conversion are typically neuter (e.g. Köpcke 1982: 74, Mills 1986: 30,

Steinmetz 1986: 198, Flämig 1991: 453, Russ 1994: 201, Hickey 2000: 630, Hoberg 2004: 90f., Menzel 2004: 63, Chan 2005: 55ff., Duden Grammatik 2009: 159, Engel 2009: 282, Eisenberg 2013: 134). These include nominalised infinitives (e.g. *das Laufen*), adjectives⁴ (e.g. *das Schwarz*), adverbs (e.g. *das Hier und Jetzt*), pronouns (e.g. *das Ich*), conjunctions (e.g. *das Wenn und Aber*), prepositions (e.g. *das Auf und Ab*), interjections (e.g. *das Hallo*) and whole verb phrases (e.g. *das Inkrafttreten*).

1.3.1.4 Compounds

A further morphological process which is relevant to gender assignment in German is compounding. The gender of a German compound is in the majority of cases determined by its rightmost element (e.g. Eisenberg 1999: 218, Donalies 2002: 57, Neef 2009: 389, Fleischer & Barz 2012: 85). For instance, *Tierarzt* is masculine because its rightmost element, *Arzt*, is masculine. As a result, a morphological gender assignment principle termed the *Letzt-Glied-Prinzip* or Last Member Principle (LMP) has been proposed, which states that, just as affixed nouns' gender is determined by their final suffix, nominal compounds assume the gender of their final constituent noun or "last member" (Köpcke & Zubin 1983: 9, 1984: 44f., 1997: 28).

However, this principle does not appear to be categorical, since there seem to be a small number of exceptions, for instance, *das Rückgrat* (despite *der Grat*), *die Wehmut* (despite *der Mut*) and *das schöne Ingolstadt* (despite *die Stadt*) (Zubin & Köpcke 1984: 47, Durrell 2002: 10, Chan 2005: 50). The Last Member Principle will be discussed further in 3.3.3 and 3.3.4.

1.3.1.5 Reductions

An additional morphological gender assignment principle which has been proposed relates to nouns which have been formed by means of reduction. These include clippings (e.g. *Uni* < *Universität*), initialisms (e.g. *SPD* < *Sozialdemokratische Partei Deutschlands*), acronyms (e.g. *FAZ* < *Frankfurter Allgemeine Zeitung*), multi-segmental reductions (e.g. *BAFög* < *Bundesausbildungsförderungsgesetz*, *KaDeWe* < *Kaufhaus des Westens*), and blends (e.g. *Kurlaub* < *Kur+Urlaub*) (e.g. Wegera 1997: 50, Donalies 2002: 144, Hoberg 2004: 91, Duden Grammatik 2009: 159).

On the basis of such nouns, a general gender assignment principle has been proposed, stating that the gender of a shortened form is determined by the gender of the corresponding full form (e.g. Wegera

⁴ This applies only to uninflected adjectives which are able to be used as predicates. Nominalised inflected adjectives, i.e. adjectival nouns, are generally considered a syntactic phenomenon rather than the product of morphological conversion (e.g. Eisenberg 1999: 282, Donalies 2002: 130, Hoberg 2004: 89, Ronneberger-Sibold 2004: 1275).

1997: 50, Hoberg 2004: 91, Chan 2005: 59). In cases where the corresponding full form consists of a single noun or compound, e.g. *Universität* or *Bundesausbildungsförderungsgesetz*, this is relatively straightforward. However, in cases where the corresponding full form is an NP – e.g. *Sozialdemokratische Partei Deutschlands*, *Kaufhaus des Westens*, *Frankfurter Allgemeine Zeitung* – the principle must be modified to specify that the gender of the shortened form is determined by the head of the NP, which can appear phrase initially, medially or finally (e.g. Wegera 1997: 50, Hoberg 2004: 91, Chan 2005: 59). Such structures are discussed further in 3.3.3.

This gender assignment principle is generally considered to be categorical, however, a small number of apparent exceptions have been identified, e.g. *das Foto* < *die Fotografie*, *das Kino* < *der Kinematograph* (e.g. Heringer 1995: 214, Wegera 1997: 50, Hoberg 2004: 91, Duden Grammatik 2009: 160). Nevertheless, Heringer (1995: 214) argues that, synchronically, these nouns are unlikely to be analysed as shortened forms, since the corresponding full forms are seldom used.

1.3.1.6 Inflection

Finally, it has also been proposed that certain inflectional properties of a German noun may too influence the gender it is assigned. Zubin & Köpcke (1981: 443) and Köpcke (1982: 79f.), for instance, claim that native German speakers may use a noun's plural inflection to determine its gender. They state that this is especially likely for nouns which occur frequently in the plural, such as *Perlen* 'pearls'. Accordingly, they propose a number of gender assignment principles based on plural morphology, shown in (1).

- (1) Nouns forming their plural with (UMLAUT) + *-er* = M/N
 UMLAUT + *-e* = M/F
-e = M/N
-(e)n = F
-s = M/N

Corbett (1991: 49) states that this approach, i.e. the use of inflectional class to determine gender, appears 'promising' for German, based on its demonstrated success for Russian (Corbett 1991: 34-43). However, not only have some linguists expressed scepticism about the role of inflection in Russian gender assignment (e.g. Galbreath 2010), Thornton (2001) argues that the directionality of the relationship between gender and inflection may well be language specific, meaning that what is true for Russian, for example, is not necessarily true for German or any other language. Her claim is based on the evidence given in Dressler & Thornton (1996) and Thornton (2001) which demonstrates that, in Italian, it is gender that determines inflection, and not inflection that determines gender.

Indeed, while it is widely accepted that there is certainly a correlation between inflectional morphology and gender in German (e.g. Salmons 1993: 414, Zifonun et al. 1997: 29f., Hickey 2000: 638, Weinrich 2007: 342, Thieroff & Vogel 2012: 40f. etc.), many linguists consider it more likely that inflectional class is determined by gender in German, rather than vice versa (e.g. Wurzel 1984, 1994; Salmons 1993: 420f.; Aronoff 1994: 74; Bittner 1994, 2000; Wegener 1995, 2000; Eisenberg 1999).

Bittner (2000), for instance, provides a number of reasons why this is more likely to be the case. In terms of singular morphology, Bittner (2000: 6) argues that a relationship whereby gender determines inflection is evident in German since the correlation between gender and case marking is so strong. As shown below in (2), all feminines are always completely unmarked for case, masculines and neuters take *-(e)s* in the genitive singular and a small group of masculines take *-(e)n* in the oblique cases.

(2)

Gender	Nominative Singular	Accusative Singular	Genitive Singular	Dative Singular
Feminine	-∅	-∅	-∅	-∅
Masculine, Neuter	-∅	-∅	-(e)s	-∅ ⁵
Masculine (weak)	-∅	-(e)n	-(e)n ⁶	-(e)n

Bittner (2000: 6) states that, given such a strong correlation, ‘gender is plainly the dominant criterion’ with regard to establishing case inflection on singular nouns, i.e. there is no factor other than gender that could so convincingly determine singular inflection. This view is also shared by linguists such as Neef (1998: 225f.) and Salmons (1993: 414), and, furthermore, by Köpcke (1982: 79f.), who although advocates the position that plural inflectional morphology can determine gender, claims that singular inflectional morphology does not affect gender assignment, since it is instead gender which determines singular inflection.

In terms of plural morphology, Bittner (2000: 8f.) argues that gender must determine inflection since the alternative is problematic. An analysis in which plural inflection determines gender relies on the premise that plural inflection is primarily determined by the phonological structure of a noun instead of its gender. However, there are three main problems with this view.

⁵ A small number of masculine and neuter nouns adopt the inflectional suffix *-e* in the dative singular. This is generally seen as archaic and largely only persists in set phrases, such as *im Sinne von* ‘in the sense of’ and *im Grunde genommen* ‘basically’ (Duden Grammatik 2009: 206f.).

⁶ There are around twelve weak masculine nouns, all of which refer to inanimates, that bear the suffix *-ns* rather than *-(e)n* in the genitive singular (see Thieroff and Vogel 2012: 45). Some grammars purport that the neuter noun *Herz* ‘heart’ also belongs to this class of nouns, since it (sometimes) bears the suffix *-en* in the dative and *-ens* in the genitive. However, unlike weak masculine nouns, it cannot take the suffix *-en* in the accusative (Duden Grammatik 2009: 218).

Firstly, referring to the phonological structure of a noun alone is not sufficient to capture important generalisations about many plural formation tendencies; it is necessary to also refer to a noun's gender in order to do this (Bittner 2000: 8). For instance, the rule "nouns ending in *-el*, *-er*, *-en* and *-lein* have a zero plural" applies only to masculine and neuter nouns, and not to feminine nouns such as *Regel* 'rule' or *Feder* 'feather', which take an *-(e)n* plural. Likewise, the rule "nouns ending in a consonant or diphthong take *-e* in the plural" also only applies to masculines and neuters. Feminines with this phonological structure take *-(e)n* in the plural, e.g. *Zeit* 'time', *Frau* 'woman'. It is therefore necessary to refer to a noun's gender when stating plural formation rules.

Secondly, Bittner (2000: 9) argues that expressing inflectional rules in terms of the phonological structure of a noun is inefficient. This is because many different phonological inputs yield same output, as shown in (3).

(3) Nouns ending in $\left. \begin{array}{l} \text{a stressed vowel} \\ \text{a consonant} \\ \text{a diphthong} \\ \text{a schwa} \\ \text{-er} \\ \text{-el} \end{array} \right\}$ (which are also feminine) form their plural in *-(e)n*

This rule could be simplified considerably if reformulated solely in terms of gender, i.e. by stating that the vast majority of feminine nouns take an *-(e)n* suffix in the plural, instead of listing all possible phonological configurations of these nouns. Such simplification appeals to the principles of economy and parsimony.

Thirdly, stating that gender determines singular inflection but that phonology determines plural inflection (as Köpcke 1982: 79f. does) is to say that singular and plural morphology constitute 'two separate and entirely independent domains within the inflectional system' (Bittner 2000: 6). Splitting singular and plural morphology in this way ignores important generalisations about how nouns of the same gender inflect in both the singular and the plural. As shown in (4), feminine nouns on the one hand, and masculine and neuter nouns on the other, tend to inflect in the same way as each other in the singular and in the plural.⁷ Similarly, weak masculine nouns are set apart from other masculines in both the singular (displaying suffix *-(e)n* in oblique cases) and in the plural (taking suffix *-(e)n*).

⁷ While not all nouns belong to these two classes, there is evidence to show that these paradigm types are both quantitatively the largest and the most productive for feminine and for masculine and neuter nouns (see Bittner 2000: 11ff.).

(4)

Gender	Singular				Plural	
	NOM	ACC	GEN	DAT	NOM.ACC.GEN	DAT
Feminine	-∅	-∅	-∅	-∅	-(e)n	-(e)n
Masculine, Neuter	-∅	-∅	-(e)s	-∅	-e	-(e)n

In order to overcome the problems associated with a split system in which gender determines singular inflection and phonology determines plural inflection, Bittner (2000: 13ff.) proposes a unified account of inflection in which gender, specifically the feature $[\pm FEM]$, determines both singular and plural inflection. In order to explain any exceptions to the paradigms in (4), Bittner (2000: 13ff.) formulates a small number of blocking rules which refer to more specific features such as the presence of certain derivational suffixes or certain phonological endings. Such an analysis is, as established above, both simpler and more efficient than the alternative.

In sum, a number of reasons have been presented which demonstrate that a relationship whereby gender is used to determine inflection is theoretically more favourable than the reverse scenario. Inflectional morphology is thus unlikely to form part of the German gender assignment system. This is in line with the majority of accounts of German gender assignment, which do not incorporate principles based on inflectional behaviour, e.g. Spitz (1965), Heringer (1995), Wegener (1995), Chan (2005) etc. Consequently, the idea that inflectional morphology plays a role in German gender assignment will not be pursued further in this study.

1.3.1.7 Evidence for morphological principles

As established above, various patterns of correlation between gender and morphological features have been identified. Furthermore, there is experimental evidence to suggest that native German speakers make use of these regularities when assigning gender to nouns.

Both Carstensen (1980) and Hohlfeld (2006), for instance, conducted experiments in which native German speakers were required to assign genders to pseudo-nouns, each of which possessed a native or loan derivational suffix. Results from both experiments showed that, in a significant majority of cases, the participants selected the gender that is predicted by the relevant morphological gender assignment principle, suggesting that native speakers make use of derivational affixes when assigning gender to nouns. Similar results have also been found in parallel experiments for Greek (e.g. Mastropavlou & Tsimpli 2011) and Dutch (e.g. Zekhnini & Hulstijn 1995).

Wegener (2000: 533) also finds evidence that speakers take advantage of morphological regularities in her study of language acquisition. In an experiment in which children assigned genders to pseudo-

nouns with native suffixes, it was shown that the participants assigned gender to the nouns in accordance with the morphological principles established, producing statistically significant results.

In summary, it can be concluded that morphological principles, which may relate to derivation, conversion or compounding but not to inflection, seem to form part of the German gender assignment system.

1.3.2 Phonological principles

It is widely held that the phonological shape of a German noun can also influence the gender it is assigned (e.g. Altmann & Raettig 1973, Köpcke 1982, Köpcke & Zubin 1983, Mills 1986, Corbett 1991, Wegener 1995, Hickey 2000). Various phonological features of nouns have been claimed to be relevant to gender assignment, resulting in the formation of numerous phonological gender assignment principles. As noted by e.g. Schmid (2002: 104), the majority of these principles are stochastic rather than deterministic in nature, i.e. are based on tendencies and not categorical rules.

1.3.2.1 Word ending

In an early study by Altmann & Raettig (1973), it was found that certain phonological endings of nouns have a statistically significant association or dissociation with certain genders. For example, it was found that nouns ending in *-ich* and nouns ending in *-ang* are strongly associated with masculine, and nouns ending in *-b* are strongly disassociated with feminine (Altmann & Raettig 1973: 302).

Since then, a variety of other principles relating to the phonological properties of the ending of a noun have been suggested. One such set of principles involve nouns with so-called “pseudo-suffixes”. Pseudo-suffixes are phonological endings which appear to be identical to a suffix, yet, synchronically at least, do not constitute morphemes (e.g. Eisenberg 1999: 203). Examples of pseudo-suffixes in German include *-el* in e.g. *Himmel* (as opposed to the suffix *-el* in e.g. *Ärmel*), *-er* in e.g. *Hammer* (as opposed to the suffix *-er* in e.g. *Lehrer*), *-en* in e.g. *Boden* (as opposed to the suffix *-en* in e.g. *Schwimmen*), and *-e* in e.g. *Biene* (as opposed to the suffix *-e* in e.g. *Höhe*).

While there exist nouns of all genders with these pseudo-suffixes, correlations between certain pseudo-suffixes and certain genders have been reported, and gender assignment principles have been formulated accordingly. Based on calculations made by Augst (1979), Wegener (1995: 75) states that 59% of nouns with the pseudo-suffix *-el*, 83% of nouns with the pseudo-suffix *-en*, and 71% of nouns with the pseudo-suffix *-er* are masculine, and 90% of nouns with the pseudo-suffix *-e* are feminine. This has led to the formation of the phonological gender assignment rules: *-el, -er, -en* → masculine and *-e* → feminine.

Other phonological gender assignment principles relating to the ending of a noun include, for example, Nelson's (1998: 214f.) proposal that nouns ending in a stressed vowel followed by /m/ or /l/ are neuter, e.g. *das Problem, das Profil*. This principle illustrates the relevance of stress patterns to gender assignment, as highlighted by e.g. Nelson (1998) and Hohlfeld (2006). For instance, Nelson (1998: 217) makes a distinction between stressed final *-on* which is claimed to be associated with masculine, e.g. *der Balkon*, and unstressed final *-on* which is claimed to be associated with neuter, e.g. *das Stadion*. He also distinguishes between stressed final *-iv*, which is claimed to be associated with neuter, e.g. *das Archiv*, and unstressed final *-iv*, which is claimed to be associated with masculine e.g. *der Imperativ*.

1.3.2.2 Other phonological factors

It is not only word-final phonological features which are claimed to be relevant to gender assignment. In his study of the gender of monosyllables in German, Köpcke (1982) proposes a number of gender assignment principles relating to various phonological properties of monosyllabic nouns, i.e. not only their ending. Köpcke (1982) categorises his proposed principles into five main groups, each of which will be outlined below.

The first category of phonological gender assignment principles consists of those which relate to the properties of the onset of a monosyllable. These include the principle stating that monosyllables with the onset /kn/ are masculine, e.g. *der Knopf*, which, according to Köpcke (1982: 89), holds for 14 out of the 15 cases listed in the Duden, and the principle stating that monosyllables beginning with /ʃ/ followed by a consonant are also masculine, e.g. *der Schrott*, which is said to hold for 144 out of the 167 cases in the Duden (Köpcke 1982: 89f.).

The second category of principles involves the properties of the nucleus of a monosyllable. For instance, Köpcke (1982: 95) states that monosyllabic nouns whose nucleus consists of a long vowel are more likely to be masculine or neuter rather than feminine, e.g. *das Ohr, der Tag*.

The third category consists of principles relating to the coda. For example, Köpcke (1982: 98, 102) states that monosyllables ending in a consonant followed by /s/ have a tendency to be masculine, e.g. *der Lachs*, and those ending in a non-sibilant fricative, i.e. /f/ or /x/, followed by /t/ have a tendency to be feminine, e.g. *die Frucht* (Köpcke 1982: 98).

The fourth category consists of so-called "stand-by" rules, which relate to a combination of the onset and nucleus or the nucleus and coda (i.e. the rhyme). An example of this is the principle stating that

monosyllables with a long, high vowel followed by /r/, e.g. *Tür*, are usually feminine (Köpcke 1982: 103).

The fifth category of phonological gender assignment principles are those based on the syllable structure of a monosyllabic noun. For instance, Köpcke (1982: 85) formulates a principle stating that nouns with the structure VCCC are either feminine or neuter, i.e. not masculine, e.g. *die Angst*. Additionally, he states that nouns with the structure CCVCC are masculine, e.g. *der Brand*, on the basis of the calculation that this is true for 111 out of the 148 cases listed in the Duden (Köpcke 1982: 84).

A final phonological gender assignment principle which is relevant to monosyllabic nouns is suggested by e.g. Wegener (1995: 87), who, based on the fact that 64% of the 1466 monosyllables investigated by Köpcke (1982) are masculine, proposes a general phonological gender assignment principle that monosyllables have a tendency to be masculine.

1.3.2.3 Evidence for phonological principles

As stated above, many of the phonological gender assignment principles that have been proposed are based on statistical tendencies. However, results from a number of studies indicate that such principles are also productive and psychologically real for native German speakers.

For instance, Köpcke & Zubin (1983) conducted an experiment in which participants had to assign genders to pseudo-nouns with specific phonological features. It was found that in most cases, the participants selected the gender predicted by the relevant phonological gender assignment principle. Mills (1986) and Wegener (1995) repeated this experiment with different participants and found largely similar results. These studies therefore suggest that native speakers make use of phonological principles when assigning gender to pseudo-nouns. Similar results have also been found in parallel studies for French (e.g. Tucker et al. 1968, 1977; Karmiloff-Smith 1979), Dutch (e.g. Zekhnini & Hulstijn 1995) and Spanish (e.g. Sadek et al. 1975).

Schiller et al. (2003) also conducted an experiment investigating the influence of phonological factors on gender assignment. In their experiment, native German speakers were required to specify the gender of various monomorphemic nouns, and their reaction times were measured. It was found that the participants' reaction times were significantly faster for nouns with a phonological feature relevant to gender, suggesting that gender assignment is influenced by phonological features.

Results from studies on child language acquisition also seem to confirm the effect of phonological features on gender assignment. In a study by Szagun et al. (2007) on the acquisition of gender by German children, it was found that errors were more frequent for nouns which did not conform to

phonological principles, thereby indicating the early use of phonological regularities in the acquisition of gender. Mills (1986) also reports early use of phonological regularities by children to determine noun gender, and Wegener (2000: 533) confirms this in an experiment using monosyllabic and polysyllabic pseudo-nouns, in which children assigned gender to the nouns in accordance with a variety of phonological principles. Similar results have also been found for e.g. French (e.g. Karmiloff-Smith 1979) and Spanish (e.g. Sadek et al. 1975).

In summary, it can be concluded that phonological principles, which may relate to various phonological features of a noun, seem to form part of the German gender assignment system.

1.3.3 Semantic principles

Various gender assignment principles have been proposed in the literature which relate to the semantic features of a noun. A number of these will be outlined below.

1.3.3.1 The Natural Gender Principle (NGP)

One of the most oft-cited semantic principles for gender assignment is what is commonly referred to as the *natürliche Geschlechtsprinzip* 'natural gender principle' (NGP) (e.g. Mills 1986: 16; Wegener 1995: 71; Köpcke & Zubin 1996: 28, 1997: 479; Menzel 2004: 60f., Chan 2005: 41). The NGP states that a noun denoting an animate will be assigned the grammatical gender which corresponds to the "natural gender" of the referent. For animals, this is their biological sex, and for humans, this is their societal gender identity. Therefore, according to the NGP, nouns denoting men and male animals are masculine, e.g. *der Mann* 'man', *der Hahn* 'rooster', and nouns denoting women and female animals are feminine, e.g. *die Frau* 'woman', *die Henne* 'hen' (e.g. Zubin & Köpcke 1981: 445, Köpcke 1982: 75f., Mills 1986: 27, Flämig 1991: 451, Köpcke & Zubin 1996: 479-484, Hoberg 2004: 99f., Weinrich 2007: 331-336, Eisenberg 2013: 137).

Additionally, the NGP states that nouns denoting young humans and young animals are neuter, e.g. *das Kind* 'child', *das Küken* 'chick' (e.g. Zubin & Köpcke 1981: 444f., Köpcke 1982: 77, Flämig 1991: 456, Eisenberg 1999: 156, Hickey 2000: 630, Hoberg 2004: 102, Chan 2005: 296, Weinrich 2007: 336). This is presumably because such animates are perceived in a sense as sexless and genderless due to their biological immaturity (see e.g. Köpcke & Zubin 1996: 483).

As noted by e.g. Chan (2005: 41), for nouns denoting people, exceptions to the NGP can generally be explained by the referent of the noun deviating somehow from the normative ideas of the societal gender categories. For instance, in terms of nouns referring to men, exceptions to the NGP are mostly

feminine nouns denoting pejorative terms for men 'who traditionally have been regarded in German culture as lacking central characteristics of culturally-defined masculinity', such as *die Tunte* and *die Schwuchtel* 'gay man [pej.]' or *die Memme* 'coward' (Zubin & Köpcke 1981: 445).

In terms of nouns referring to women, exceptions to the NGP generally constitute neuter nouns denoting either young women (as above), e.g. *das Mädchen*, *das Gör*, or pejorative terms for women based on factors such as their culturally-determined 'sexual undesirability' because of old age, e.g. *das Weib*, *das Reff* 'old woman [pej.]', or perceived behaviour, e.g. *das Luder*, *das Mensch* 'wanton woman [pej.]' (Zubin & Köpcke 1981: 445, Köpcke & Zubin 1996: 483; Mills 1986: 16).

In terms of nouns referring to animals, there appear to be quite a number of nouns which are not assigned gender according to the NGP. This is because the NGP is largely limited to domestic animals, such as *der Hengst* 'stallion' – *die Stute* 'mare' – *das Fohlen* 'foal' or *der Bulle/Stier* 'bull' – *die Kuh* 'cow' – *das Kalb* 'calf' (e.g. Zubin & Köpcke 1981: 444, 1986: 154ff.; Köpcke 1982: 76f.), presumably because their sexed terms are traditionally of cultural importance (see e.g. Zubin & Köpcke 1986: 153).

The grammatical gender of many nouns denoting animals, however, is not based on their biological sex. It is claimed, for instance, that terms for domestic animals which are unspecified for sex are usually neuter, e.g. *das Rind* 'cow', *das Pferd* 'horse', *das Schwein* 'pig' (Zubin & Köpcke 1981: 444; 1986: 153). Animals whose sex is simply not of cultural relevance, on the other hand, can be of any of the three genders, e.g. *die Spinne* 'spider', *der Koala*, *das Krokodil* 'crocodile'. The grammatical gender of such nouns is claimed to sometimes be based on other semantic factors. Köpcke & Zubin (1996: 484), for example, claim that classes of animals which are larger and more human-like, such as mammals, have a tendency to be masculine, e.g. *der Affe* 'monkey', *der Elephant*, and classes of animals which are smaller and less human-like, such as insects, have a tendency to be feminine, e.g. *die Fliege* 'fly', *die Laus* 'louse'.

1.3.3.2 Other semantic classes

Other semantic principles for gender assignment relate to the association of certain semantic classes with particular genders. It is claimed that a noun has a tendency to be assigned a particular gender if it is a hyponym of a certain category. For example, masculine is generally assigned to currencies, e.g. *der Euro*, *der Dollar*, and alcoholic drinks, e.g. *der Wein*, *der Wodka*, feminine is generally assigned to trees, e.g. *die Eiche*, *die Kiefer*, and fruit, e.g. *die Mango*, *die Kiwi*, and neuter is generally assigned to metals, e.g. *das Gold*, *das Aluminium*, and scientific units, e.g. *das Volt*, *das Watt* (e.g. Zubin & Köpcke 1981: 444, 1984: 44, 1996: 479f., Köpcke 1982: 72ff., Flämig 1991: 452f., Hickey 2000: 629f., Hoberg

2004: 106, Menzel 2004: 63, Thomoglou 2004: 34, Chan 2005: 45, 94f., Engel 2009: 272, Weinrich 2007: 328, Duden Grammatik 2009: 158f., Eisenberg 2013: 138).

A principle termed the *Leitwortprinzip* (LWP) dictates that nouns in these semantic categories are assigned the same gender as that of their hypernym (Wegener 1995: 72). That is to say, metals, for example, are claimed to be assigned neuter because *Metall* is neuter, and dances assigned masculine due to the masculine gender of *Tanz* 'dance' (Wegener 1995: 72). However, as evident from the examples above, this principle does not seem to hold for many semantic classes. *Obst* 'fruit', for instance, is neuter, yet hyponyms of *Obst* are usually feminine, *Baum* 'tree' is masculine, yet hyponyms of *Baum* are usually feminine, and *alkoholisches Getränk* 'alcoholic drink' is neuter, yet hyponyms of this category are usually masculine. It thus seems that the gender associated with hyponyms of a particular semantic category is not linked to the gender of the category head.

What the LWP does appear to account for, however, is the fact that hyponyms of certain relatively narrow semantic categories do seem to be assigned the same gender as that of their hypernym, e.g. *der Wein* > *der Bordeaux*, *der Rioja*; *die Zigarette* > *die Lord*, *die Marlboro*; *das Hotel* > *das Hilton*, *das Ritz*. It seems, therefore, that the validity of the LWP could be dependent on the level of abstraction of the hypernym, which will be explained in detail below.

1.3.3.3 Levels of abstraction

A final group of semantic gender assignment principles relate to the level of abstraction at which the referent of a noun is organised. It is claimed that concepts are organised at different levels of a hierarchical structure according to their degree of specificity (e.g. Rosch 1977). The three main levels are: superordinate > basic level > subordinate, whereby a superordinate term is the least specific and a subordinate term is the most specific. An example is given in Figure 1.

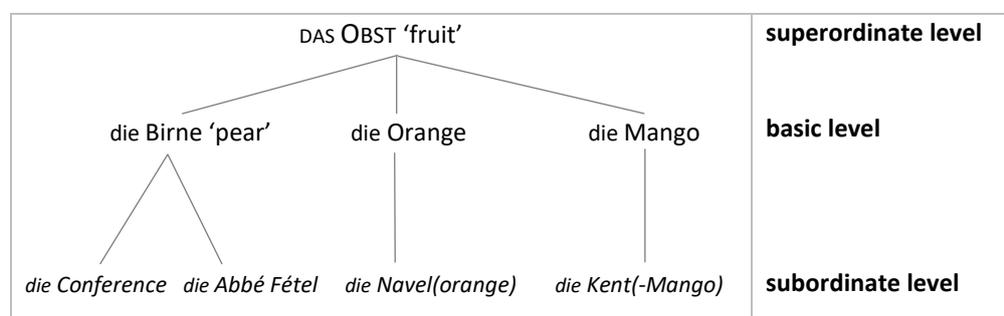


Figure 1: An (incomplete) taxonomy hierarchy of *Obst*

All levels of the hierarchy are said to be defined in terms of the basic level, i.e. superordinates are organised on the level of abstraction above the basic level, and subordinates on the level below the basic level. It is claimed that terms on the basic level can be determined via various experimental methods, since this is the level at which things are argued to be cognitively organised (e.g. Rosch et al. 1976, Rosch 1977). See Section 3.2.2.2 for a fuller discussion on this topic.

In terms of gender assignment, it is claimed that nouns on the superordinate level are usually neuter in German, e.g. *das Obst* 'fruit', *das Tier* 'animal', *das Element* (Zubin & Köpcke 1986; Mills 1986: 27; Steinmetz 1986: 192, 2006: 1424; Köpcke & Zubin 1997: 35-38, 2005: 152; Eisenberg 1999: 156; Rice 2006: 1398; Duden Grammatik 2009: 158).

Basic level terms, by contrast, can be of any gender. As mentioned above, basic level terms may be assigned a particular gender if they are hyponyms of a certain category, e.g. basic level hyponyms of *Obst* are usually feminine. The gender they are assigned is not necessarily the same as that of their superordinate (contrary to the LWP).

Finally, subordinate terms are claimed to typically be of the same gender as their hypernymous basic level term (e.g. Köpcke & Zubin 2005a: 98, 2005b: 152). Often, in German, a subordinate has the same gender as its corresponding basic level term because it is a compound with the basic level term as its head, e.g. *die Orange* > *die Navelorange*, but this is not always the case. It is at this level of abstraction, then, that the LWP gains plausibility. This observation is in line with Thornton's (2009) Basic Level Hyperonym Constraint, formulated on the basis of evidence from Italian and Hausa, which states that a hyponym can only inherit the gender of a hypernym if the hypernym is a basic level (and not superordinate) term.

Köpcke & Zubin (2005a: 111f.), however, highlight that not all subordinates inherit the gender of the hypernymous basic level term. For instance, they state that hyponyms of the neuter basic level term *Pferd* 'horse' are usually masculine, e.g. *der Holsteiner*, *der Braune*, *der Appaloosa*, and hyponyms denoting car types are usually masculine despite the fact that the hypernymous basic level term could justifiably be *das Auto* and not *der Wagen* (see Köpcke & Zubin 2005a: 107, 112). The LWP is therefore unable to replace all semantic gender assignment principles at the subordinate level.

A summary of the semantic gender assignment principles discussed in this section is presented in Figure 2.

NGP	men and male animals → m. women and female animals → f. derogatory terms for men → f. derogatory terms for women → n.	<i>der Onkel</i> <i>die Schwester</i> <i>die Tunte</i> <i>das Mensch</i>
superordinate level	superordinates → n.	<i>das Obst</i>
basic level	alcoholic drinks → m. currencies → m. mammals → m. fruit → f. trees → f. insects → f. scientific units → n. metals → n.	<i>der Wodka</i> <i>der Euro</i> <i>der Elephant</i> <i>die Mango</i> <i>die Kiefer</i> <i>die Fliege</i> <i>das Volt</i> <i>das Gold</i>
subordinate level	wines → m. horses → m. cars → m. cigarettes → f. hotels → n.	<i>der Rioja</i> <i>der Appaloosa</i> <i>der Fiat</i> <i>die Lord</i> <i>das Hilton</i>

Figure 2: Summary of semantic gender assignment principles

1.3.3.4 Evidence for semantic principles

Not only have various patterns of correlation been observed regarding a noun's semantic features and its gender, but there is also experimental evidence which suggests that native speakers use semantic principles in order to assign gender to nouns.

Schiller et al. (2003), for instance, conducted a study in which they investigated the influence of semantic factors on gender assignment. In their experiment, native German speakers were required to specify the gender of various monomorphemic nouns, some of which had gender-relevant semantic features and some of which did not, and their reaction times and ERPs were measured. The ERP data and reaction times recorded demonstrate that participants were faster in selecting the gender of nouns which had a relevant semantic feature than for those which did not. Consequently, Schiller et al. (2003) conclude that native speakers are influenced by semantic factors when assigning gender.

A further experiment investigating the use of semantic gender assignment principles by native German speakers was conducted by Schwichtenberg & Schiller (2004). Their participants were presented with a semantic category and two formally-similar pseudo-nouns of different genders, and were asked to select the noun that was more likely to belong to the category, e.g. type of fruit: *der Jast* or *die Mauch*? The results showed that in a significant majority of cases, participants selected the pseudo-noun of the gender which is claimed to be associated with the semantic category. Accordingly, they conclude that semantic regularities are part of the German gender assignment system.

Lang (1976) also conducted an experiment using pseudo-nouns, in which participants were required to assign genders to a number of pseudo-nouns with an unspecified meaning. Subsequently, the

participants were presented with a definition for each pseudo-noun, and were again asked to assign genders to each. In many cases, the distribution of results obtained with and without the definition was significantly different, demonstrating that the semantic features of the pseudo-nouns affected the participants' choice of gender.

Additionally, a number of studies on loanwords indicate that semantics is key to gender assignment. Mills (1986:50f.), for example, states that evidence from loans shows that semantic features influence gender assignment, given that various loanwords receive the gender that is predicted by proposed semantic principles. Carstensen (1980: 23) and Fischer (2005: 287ff.), too, illustrate the influence of a loan's semantic features on its gender, both finding that gender variation among speakers occurs more when the meaning of a loan is unknown, and that gender selection is more likely to be unanimous when the meaning is known, thus showing that gender assignment is affected by knowledge of a loanword's meaning.

Finally, Wegener (2000: 514f., 531f.) claims that evidence from child language acquisition demonstrates that the NGP influences gender assignment from an early age, reporting semantically-motivated gender errors even with high-frequency nouns, e.g. **die_{FEM.SG} Mädchen_{NEUT}* 'girl' and **der_{MASC} Kind_{NEUT}* 'child' (when referring to a boy). Wegener (2000: 514f.) claims that if gender were simply memorised rather than derived, it would be expected that only low-frequency nouns would cause problems.

In summary, it has been shown that there are numerous morphological, phonological and semantic gender assignment principles which can be used to determine the gender of German nouns. Additionally, experimental evidence has been presented which suggests that native German speakers actually make use of these principles in their assignment of gender to nouns. In the next section, it will be considered how these various principles interact with each other.

1.4 Interaction of gender assignment principles

One aspect of gender assignment that has received comparatively little attention in the literature is the interaction of gender assignment principles. Often, principles are simply presented as lists of disconnected rules, as in traditional reference grammars (e.g. Flämig 1991, Weinrich 2007, Engel 2009), without any mention of their possible interaction. That is to say, it is unclear what happens when a noun is subject to multiple principles which conflict in their assignment of gender. This is key, since a comprehensive theory of German gender assignment requires a systematic method of

resolving conflicts between principles (e.g. Steinmetz 1985: 10, Menzel 2004: 69, Rice 2006). The main proposals for principle interaction that have been suggested will be discussed below.

1.4.1 Semantics » form

The first proposal to be discussed states that all gender assignment systems have a “semantic core” (Aksenov 1984: 17f., Corbett 1991: 307) or “semantic base” (Lang 1976). This means that, in all languages, ‘the formal gender assignment rules [...] are dominated by the semantic gender assignment rules’ (Corbett & Fraser 2000: 321), i.e. when there is a conflict between gender assignment principles of different types, ‘the semantic rule takes precedence’ (Corbett 1991: 66).

However, Rice (2004: 1412) argues that this approach is ‘untenable’ for German, given the large number of counterexamples that exist. For instance, he cites the three nouns *die Pflanze*, *die Waffe* and *die Frucht*, stating that their gender is determined by the dominance of the formal principles “nouns with a final schwa are feminine” (e.g. Altmann & Raettig 1973: 302, Mills 1986: 33, Steinmetz 1986: 192, Steinmetz & Rice 1989: 166, Wegener 1995, Köpcke & Zubin 1996: 476, Nelson 1998: 218, Hoberg 2004: 87, Rice 2006: 1396, Eisenberg 2013: 134) and “monosyllabic nouns ending in a non-sibilant fricative [f]/[x]/[ç] followed by [t] are feminine” (e.g. Zubin & Köpcke 1981: 440, 1984: 44, Köpcke 1982: 98, Mills 1986: 33, Köpcke & Zubin 1997, Hoberg 2004: 98, Menzel 2004: 68, Duden Grammatik 2009: 166, Eisenberg 2013: 135) over the semantic principle “nouns denoting a superordinate are neuter” (e.g. Zubin & Köpcke 1986, Mills 1986: 27, Steinmetz 1986: 192, 2006: 1424, Eisenberg 1999: 156, Rice 2006: 1398, Duden Grammatik 2009: 158).⁸

Other counterexamples include: *die Gymnastik*, whose gender is determined by the dominance of the formal principle “nouns in *-ik* are feminine” (e.g. Altmann & Raettig 1973: 302, Zifonun et al. 1997: 33, Nelson 1998: 218, Hoberg 2004: 87, Menzel 2004: 66, Thomoglou 2004: 41, Weinrich 2007: 327, Duden Grammatik 2009: 165, Engel 2009: 281) over the semantic principle “nouns denoting sports and games are neuter” (e.g. Mills 1986: 27, Köpcke & Zubin 1996: 480, Chan 2005: 96); *der Wacholder*, whose gender is determined by the dominance of “nouns with the pseudo-suffix *-er* are masculine” (e.g. Steinmetz 1986: 197, Wegener 1995, Hoberg 2004: 92, Eisenberg 2013: 135) over “nouns denoting trees are feminine” (e.g. Flämig 1991: 452, Hickey 2000: 629, Menzel 2004: 63, Thomoglou 2004: 34, Chan 2005: 95, Weinrich 2007: 328, Engel 2009: 272); and *der Mohn*, whose gender is determined by the dominance of “monosyllables with the structure CVC are not feminine” (Köpcke 1982: 85) and “monosyllables containing a long vowel and a consonant in the coda are not feminine”

⁸ Naturally, the validity of this criticism is highly dependent on the validity of these gender assignment principles, a point which will be discussed in 1.4.5.3 and Chapter 2.

(Köpcke 1982: 87) over “nouns denoting flowers are feminine” (e.g. Flämig 1991: 452, Hickey 2000: 630, Menzel 2004: 63, Thomoglou 2004: 34, Chan 2005: 95, Weinrich 2007: 328).

A further problem with this approach, and indeed any approach which ranks principles according to their broad type, i.e. meaning and form, is that it does not provide a method of conflict resolution in cases where the conflicting principles are of the same type, i.e. all formal or all semantic. For example, *der Schwur* is subject to the principle “monosyllables containing a long, high vowel followed by /r/ are feminine” (Köpcke 1982: 103) as well as “monosyllables with the structure CCVC are masculine” (Köpcke 1982: 85) and “monosyllables with an onset consisting of [ʃ] followed by a consonant are masculine” (e.g. Köpcke 1982: 89f., Mills 1986: 33, Köpcke & Zubin 1997, Hoberg 2004: 98, Menzel 2004: 68, Chan 2005: 296), and *die Gemeinschaft* is subject to both “nouns in *Ge-* are neuter” (e.g. Mills 1986: 30, Hickey 2000: 631, Menzel 2004: 66, Rice 2006: 1396, Steinmetz 2006: 1424) and “nouns in *-schaft* are feminine” (e.g. Mills 1986: 30, Flämig 1991: 453, Zifonun et al. 1997: 33, Hickey 2000: 630, Hoberg 2004: 87, Menzel 2004: 66, Thomoglou 2004: 41, Weinrich 2007: 327, Duden Grammtik 2009: 165, Engel 2009: 281, Eisenberg 2013: 133), yet a meaning over form approach does not resolve these conflicts.

A key proponent of the semantics » form approach is Köpcke (1982), who claims that German gender assignment principles are ordered hierarchically, such that semantic principles dominate morphological principles, which in turn dominate phonological principles, which are themselves ordered according to the following hierarchy: coda » “stand-by” » onset » structure » nucleus (see 1.3.2). Whilst Köpcke’s (1982) model is more detailed than a broad semantics » form approach, it is nevertheless unable to account for the examples listed above, and, moreover, there are some specific problems with his model.

Firstly, Köpcke’s model is based exclusively on monosyllabic nouns, meaning that it does not incorporate any generalisations made on the basis of polysyllables (as highlighted by e.g. Chan 2005: 69), thereby excluding most derivational morphological principles and any phonological principles based on polysyllables (e.g. pseudo-suffixes). Secondly, the morphology referred to in Köpcke’s model consists only of inflectional morphology, which, as seen in 1.3.1.6, is problematic. Thirdly, as highlighted by Wegener (1995: 77) and Fischer (2005: 115), there are some cases in Köpcke’s study in which his model predicts two possible genders for a noun, where only one of these is actually correct. However, Köpcke (1982: 11) misleadingly classes these cases as successful predictions.

A final criticism of Köpcke’s model comes from Heringer (1995: 214), who considers Köpcke’s (1982: 126) argument that the neuter gender of diminutives in e.g. *-chen/-lein* should be explained by the semantic principle “nouns denoting diminutives are neuter” rather than the morphological principle

“nouns with the suffix *-chen/-lein* are neuter” (thereby justifying a semantics » form approach in cases like *Mädchen*) problematic. For Heringer, the assignment is undoubtedly due to the form, arguing that nouns with diminutive meaning are not always neuter in German anyhow (e.g. *der Vati, die Mutti* – as also noted in e.g. Dressler & Barbaresi 1994: 104). Dahl (2000: 103), too, argues that it is synchronically more appropriate to classify the principle as form-based rather than semantic.

The semantics » form approach to principle interaction thus seems unsuitable for German.

Accordingly, a less extreme version of the semantics » form approach to principle interaction has been suggested by Nessel (2003, 2005, 2006a, 2006b), who proposes the Core Semantic Override Principle (CSOP). The CSOP states that it is only semantic gender assignment principles relating to societal gender/sex (i.e. men/males = masculine and women/females = feminine) which dominate all other principles. Nessel’s claim is largely based on data from Ukrainian, but he states that the principle is universal.

However, Dahl (2000: 103ff.) identifies three main classes of exceptions based on typological data. Firstly, he states that morphological principles relating to augmentative and diminutive derivations often override the CSOP, citing the German case of animates in *-chen* and *-lein*. Secondly, he identifies cases of the CSOP being overridden by semantic principles denoting young or small animates, such as in certain Polish dialects (e.g. Corbett 1991: 100), and also in German, e.g. *das Kind, das Küken* (see 1.3.3.1).

Thirdly, there are cases in which semantic principles denoting certain kinds of animals override the CSOP, such as in Tamil, Dyirbal and Ngangikurrungurr, in which animals are not assigned gender according to their sex (Dahl 2000: 104f.). In German, this is seen with domestic animals, which are typically neuter, e.g. *das Rind, das Pferd, das Schwein*, and animals which are claimed to be assigned gender according to their relative size, e.g. *der Elephant, die Fliege* (Zubin & Köpcke 1981: 444; 1986: 153, 484) (see 1.3.3.1). Additionally, there are cases in German where the CSOP is overridden by derogatory terms, e.g. *die Memme, das Weib*, as seen in 1.3.3.1.

The CSOP therefore also seems unsuitable for German.

1.4.2 Form » semantics

Given that the semantics » form approach appears unable to account for all of the German data, the natural alternative to consider is the converse view, namely that all formal principles override all semantic principles. However, again, there are numerous counterexamples to this approach. Rice (2006: 1411), for instance, cites cases such as *Bote* ‘messenger’, whose masculine gender is

determined by the dominance of the semantic principle “nouns denoting men are masculine”⁹ (e.g. Zubin & Köpcke 1981: 445, Köpcke 1982: 75, Mills 1986: 27, Flämig 1991: 451, Köpcke & Zubin 1996: 479, Hoberg 2004: 99, Weinrich 2007: 331, Duden Grammatik 2009: 154, Eisenberg 2013: 137) over the formal principle “nouns with a final schwa are feminine” (see above). This is also the case for nouns such as *der Junge* ‘boy’.

Additional counterexamples include *der Euro*, whose gender is determined by the dominance of the semantic principle “nouns denoting currencies are masculine” (e.g. Menzel 2004: 63, Thomoglou 2004: 34, Chan 2005: 94) over the formal principle “nouns with a final (unstressed) -o are neuter” (e.g. Hickey 1999: 641, Menzel 2004: 68); *die Mango*, whose gender is determined by the dominance of “nouns denoting types of fruit are feminine” (e.g. Köpcke & Zubin 1996: 480, Duden Grammatik 2009: 158) over “nouns with a final (unstressed) -o are neuter”; and *der Tequila* and *der Wodka*, whose gender is determined by the dominance of “nouns denoting alcoholic drinks are masculine” (e.g. Zubin & Köpcke 1981: 444, Köpcke 1982: 72, Köpcke & Zubin 1996: 479, Hickey 2000: 630, Menzel 2004: 63, Thomoglou 2004: 34, Chan 2005: 94, Engel 2009: 272, Duden Grammatik 2009: 159, Eisenberg 2013: 138) over “nouns with a final -a are feminine” (e.g. Steinmetz 1986: 197, Steinmetz & Rice 1989: 166, Thomoglou 2004: 41, Durrell 2011: 8, Duden Grammatik 2009: 164, Kraiss 2014: 36).

A further problematic aspect of this approach, as mentioned above, is the fact that it cannot provide a method of conflict resolution in cases where the conflicting principles are of the same type, i.e. both formal or both semantic.

It therefore seems that the form » semantics approach is also unable to account for all of the German data.

1.4.3 Morphology » semantics » phonology

Another approach to German gender assignment that has been proposed requires the division of formal principles into those pertaining to a noun’s morphology and those pertaining to a noun’s phonology, such that morphological principles override semantic principles, which in turn override phonological principles. This approach is advocated by Wegener (1995: 88), who states that it is able to account for cases such as *Mädchen* and *Rotkäppchen* (examples of morphology » semantics), and *Knecht* and *Wicht* (examples of semantics » phonology). She adds that the morphology referred to in her model excludes inflectional features, unlike in that of Köpcke (1982).

⁹ Or “nouns denoting a generic person without specific reference to their societal gender are masculine” (e.g. Köpcke 1982: 72, Mills 1986: 27, Hoberg 2004: 103, Chan 2005: 91ff., Onysko 2007: 159).

However, there are again a number of exceptions to this approach. The aforementioned examples of *Pflanze*, *Waffe*, *Frucht*, *Wacholder* and *Mohn* (see 1.4.1) all constitute examples of phonological principles overriding semantic principles, thereby demonstrating that the morphology » semantics » phonology approach to principle interaction is also unable to account for all of the German data.

A very similar approach is taken by Heringer (1995: 214), who proposes the following hierarchy: composition » derivation » NGP » semantic category (LWP) » coda. However, as well as being unable to account for the above examples, there are some specific problems with his model. Firstly, he fails to acknowledge principles relating to conversion, and any phonological principles relating to the onset or syllable structure. Secondly, he includes only the LWP in terms of semantic categories, thereby excluding any semantic principles which do not assign gender based on the gender of the hypernym. Finally, his placement of the NGP above all other semantic principles is problematic for the reasons listed in 1.4.1.

A further criticism of the morphology » semantics » phonology approach is that, as highlighted by Corbett (1991: 33), the distinction between morphological and phonological principles is not always clear cut, thereby making their separation problematic. For instance, Wegener (1995: 73f.) lists *-ig* as in *König*, *Honig*, *Käfig*, as a derivational suffix, whereas this might more appropriately be analysed synchronically as just a phonological ending, since it is neither productive nor transparent nor obviously pattern-forming. The same applies to cases such as *-ade* in *Marmelade* and *Fassade*, *-age* as in *Garage* and *Etage*, and several other loan suffixes, including *-ik*, *-ma*, *-ur*, *-ette* and *-ett*, which may be transparent as suffixes in some nouns but unanalysable as such in others.

A final problematic aspect of this approach is the aforementioned argument that it cannot provide a method of conflict resolution in cases where the conflicting principles are of the same type, i.e. both/all morphological, both/all semantic or both/all phonological.

For all of the reasons listed above, it appears that this approach to gender assignment is also unfavourable for German.

1.4.4 Hierarchy of individual principles

Another possible arrangement of gender assignment principles is the ordering of principles individually. One such example of this is the model suggested by Spitz (1965), who proposes a set of 20 gender assignment rules which are each individually ranked with respect to one another. The rules he proposes are broad in nature, with many referring to multiple features, e.g. “nouns ending in *-o/* and *-um* are neuter” (Spitz 1965: 42), meaning that the rule set involves around 40 different features

in total. However, as with the aforementioned approaches to principle interaction, there are numerous exceptions to his model, causing linguists such as Rogers (1987: 54) to dismiss it entirely. Moreover, there is no indication as to what happens in cases where a noun is not covered by any of the 20 rules. Spitz (1965: 42) himself acknowledges that the model is 'bei weitem nicht vollständig [nowhere near complete]'.

The ordering of individual principles has also been argued to be ineffective by Rice (2006: 1413f.), who cites the case of the four nouns *Pflanze*, *Waffe*, *Gemüse* and *Gemeinde*. He states that in order for both *Pflanze* and *Waffe* to be correctly assigned feminine gender, the hierarchy "nouns with final schwa are feminine" over "nouns denoting a superordinate" is required. Additionally, in order for *Gemeinde* to be correctly assigned feminine, the hierarchy "nouns with a final schwa are feminine" over "nouns in *Ge-* are neuter" is required. However, with these hierarchies in place, *Gemüse* (neuter), which is subject to all three of these principles,¹⁰ would be incorrectly predicted to be feminine too. Examples such as these lead Rice (2006: 1414) to conclude that there is no single hierarchy of individual principles which accounts for the genders of all German nouns.

1.4.5 Optimal Gender Assignment Theory (OGAT)

A final approach to principle interaction comes in the form of Optimal Gender Assignment Theory (OGAT). Unlike the aforementioned hierarchical approaches to gender assignment, a central aspect of OGAT is the crucial non-ranking of gender assignment principles. That is to say, OGAT rejects the hierarchical ordering of individual gender assignment principles (contra e.g. Spitz 1965), as well as the hierarchical ordering of principles by type (contra e.g. Lang 1976, Köpcke 1982, Aksenov 1984, Corbett 1991, Heringer 1995, Wegener 1995). This is because OGAT considers form and meaning to be of equal importance to gender assignment. In this section, the main features of OGAT will be outlined (1.4.5.1) and the strengths of the theory will be considered (1.4.5.2). Subsequently, the weaknesses and aspects of OGAT that require further investigation will be discussed (1.4.5.3) and previous testing of OGAT in the literature will be reviewed (1.4.5.4).

1.4.5.1 Theory outline

The fundamental basis for OGAT is the work of Steinmetz (1985, 1986, 2001, 2006). Steinmetz claims that there is no hierarchical ordering of gender assignment rules but instead that their interaction is governed by the two principles of 'gender tally' and 'gender eclipse'. Gender tally refers to counting

¹⁰ There is some debate as to whether "*Ge- = n*" and "*-e = f*" should be considered two separate rules or just one rule "*Ge_e = n*" in cases such as this (see e.g. Plank 1986 and Rice 2006: 1398, and Section 3.3.3.3).

the number of times each gender is assigned to a particular noun by all of the applicable rules, and then assigning the noun the gender with the highest tally value. For example, in (5), *Gemüse* is assigned neuter, since there are two applicable rules which assign neuter and only one rule which assigns feminine.

- (5) *Gemüse* 'vegetable'
 Ge- = n
 -e = f
 superordinate = n
0m > 1f > 2n = n
 (Steinmetz 2006: 1425)

Gender eclipsis, on the other hand, comes into effect when two or more genders end up with the same gender tally value. Gender is then assigned according to a language-specific gender hierarchy in which, for German, masculine “eclipses” feminine which in turn outweighs neuter. In (6), *Pflanze* is assigned feminine on the basis of feminine being hierarchically superior to neuter according to the principle of gender eclipsis.

- (6) *Pflanze* 'plant'
 -e = f
 superordinate = n
0m > 1f > 1n = f
 (Steinmetz 2006: 1425)

The principle of gender eclipsis also means that nouns to which no rules apply are automatically assigned masculine, since masculine is at the top of the proposed gender hierarchy (Steinmetz 2006: 1425).

Using the work of Steinmetz as a foundation, Rice (2004, 2005, 2006) proposes a theory of gender assignment which is embedded in the framework of Optimality Theory (OT) called Optimal Gender Assignment Theory (OGAT), which he develops for Russian (Rice 2004, 2005), German, French, Norwegian and Dutch (Rice 2006). Central to Rice’s theory is Steinmetz’s proposal that all types of feature relevant to gender assignment are equally important, i.e. there is no hierarchy of formal and semantic gender assignment principles.

Rice (2006: 1395) realises this in OT through the ‘crucial non-ranking’ of all gender assignment constraints, regardless of whether they pertain to a noun’s form or meaning. Here, Rice uses “crucial” in the sense of Prince & Smolensky (1993/2004: 61), meaning that the constraints *cannot* dominate each other; they must be equally ranked. This is directly opposed to non-crucial non-ranking, which would mean that the ranking of the constraints is simply unimportant, i.e. the outcome would be the

same with any ranking order. Rice (2006: 1394f.) claims that this crucial non-ranking of constraints in OT is deemed possible but has not thus far been explored by Prince & Smolensky (1993/2004).¹¹

In line with standard OT convention, Rice (2004, 2005, 2006) illustrates OGAT using OT tableaux, in contrast to the gender assignment tables used by Steinmetz. The three possible genders of a noun constitute the candidates (generated on the basis of the noun as the genderless input), which are then evaluated according to a set of gender assignment constraints, which essentially replace the gender assignment rules used in Steinmetz's analyses. The gender assignment constraints are all crucially equally ranked with respect to one another and are expressed as a block of GENDER FEATURES. Rice (2006: 1405) states that in logical terms, these equally-ranked GENDER FEATURES constraints 'function as disjunctive elements of one constraint'.

When a candidate violates a constraint, it is marked with an asterisk. Once a candidate incurs more violations than another candidate, it is given a fatal violation, illustrated by an exclamation mark. A fatal violation leaves a candidate unable to surface as the favoured candidate, meaning that it is effectively eliminated from the candidate set. The candidate with the fewest violations then emerges as the optimal candidate, which is indicated by ☞ (or ☞ if the incorrect outcome is produced). This aspect of OGAT is essentially a reflection of Steinmetz's principle of gender tally through the prism of OT.

The tableau shown in (7) is the OGAT version of the gender table shown in (5). The three candidates *der*, *die* and *das Gemüse* are generated on the basis of the input *Gemüse*, and are then evaluated according to the three equally-ranked GENDER FEATURES constraints: $-E \rightarrow *M, *N$ 'nouns ending in a schwa cannot be masculine or neuter'; $GE \rightarrow *M, *F$ 'nouns with the prefix *Ge-* cannot be masculine or feminine'; and $SUP \rightarrow *M, *F$ 'nouns denoting superordinates cannot be masculine or feminine'. In (7), *das Gemüse* correctly emerges as the optimal candidate since it incurs the fewest violations within the GENDER FEATURES block.

¹¹ Other scholars who have explored the possibility of crucial equal ranking include Ní Chiosáin (1999), who demonstrates how crucial non-ranking accounts for Irish phonotactics, Crowhurst (2001), who uses crucial non-ranking to account for *um*-infixation in Toba Batak, Crowhurst & Michael (2005), who account for stress in Nanti with crucial non-ranking, Topintzi (2005a, 2005b), who uses crucial non-ranking to account for stress in Arabela, and Müller (1999) and Schmid (2001), who highlight crucial non-ranking as a possible way of accounting for optionality in OT.

(7)

<i>Gemüse</i>	GENDER FEATURES ¹²			*N	*F	*M
	-E→*M,*N	GE→*M,*F	SUP→*M,*F			
<i>der Gemüse</i>	*	*	*!			*
<i>die Gemüse</i>		*	*!		*	
<i>☞ das Gemüse</i>	*			*		

Rice (2006: 1408)¹³

It can, however, be the case that two or more candidates share the same lowest number of violations of the constraints within the GENDER FEATURES block, meaning that no one candidate can surface as the optimal on the basis of these constraints alone. When this occurs, Rice (2006) proposes that the candidates are evaluated according to a default markedness hierarchy which is ranked below the block of equally-ranked GENDER FEATURES. This aspect of the theory is based on Steinmetz's principle of gender eclipse. Rice (2006) maintains that, in German, the least marked gender is masculine and the most marked gender is neuter, the optimal gender being the one which is least marked, i.e. masculine. The main reasoning behind this ordering is the claim that masculine is the category to which the largest number of nouns belong (Rice 2006: 1406). This point will be discussed further in 1.4.5.3.

An example of a noun whose gender is determined by the default markedness hierarchy, since it cannot be determined by referring only to the block of GENDER FEATURES constraints, is *Pflanze*, as shown in (8), which is the OGAT version of the gender table in (6). *Die Pflanze* and *das Pflanze* incur only one violation each in the GENDER FEATURES block, meaning that the markedness hierarchy is required to determine the optimal candidate. Since neuter is the most marked and therefore least optimal gender, *das Pflanze* immediately incurs a fatal violation, leaving *die Pflanze* to emerge as the optimal candidate.

(8)

<i>Pflanze</i> 'plant'	GENDER FEATURES		DEFAULT HIERARCHY		
	-E→*M,*N	SUP→*M,*F	*N	*F	*M
<i>der Pflanze</i>	*	*!			*
<i>☞ die Pflanze</i>		*		*	
<i>das Pflanze</i>	*		*!		

Rice (2006: 1408)

¹² The order in which the constraints are listed in the GENDER FEATURES block is irrelevant, since the constraints are all equally ranked, i.e. the leftwards placement of -E→*M,*N should not be construed as hierarchical dominance. In line with this, fatal violation markings have been placed after the rightmost violation in the GENDER FEATURES block (unlike in Rice (2006), where they appear after individual GENDER FEATURES constraints), since all individual GENDER FEATURES constraints are essentially functioning together as one.

¹³ Rice (2006: 1405), like McCarthy (2002), does not shade cells to the right of a fatal violation, labelling this OT convention as 'redundant'.

Rice (2004, 2005, 2006) chooses OT as the framework for his proposed model of gender assignment for two main reasons. Firstly, Rice (2004: 6) considers OT a ‘natural choice’ for a framework in which to explore gender assignment, since it is ‘a theory designed specifically to mediate conflicts between violable constraints’ (Rice 2006: 1404), which is the process at the core of gender assignment. He claims that the apparent probabilistic nature of many German gender assignment principles can be explained in OT as an expected result of the resolution of conflicts between violable GENDER FEATURES constraints (Rice 2006: 1395), and that his use of OT is in part motivated by the central role of violable constraints in the theory (Rice 2006: 1397).

Secondly, Rice (2006: 1408) explains that using OT enables the theory to overcome a potentially problematic aspect of Steinmetz’s model, namely the fact that it requires counting.

It is widely accepted that many rules in generative grammar are bound by locality principles, i.e. that rules often only apply within a limited distance of a particular structure (Hayes 1995: 307, Nessel 2006a: 1375). Hayes (1995: 307) observes that ‘the principle of locality often takes the form of limiting what can be counted’. This has led to the belief that grammars are essentially non-quantitative, i.e. “grammars cannot count”, or at most weakly quantified (Guy 2011: 2195f.) (Hayes (1995: 307), for instance, hypothesises that phonological rules can count up to two). In accordance with this, Nessel (2006a: 1375) notes that languages do not seem to posit rules which involve counting linguistic units. For example, there appears to be no language which proposes a phonological rule involving, say, applying stress to the fourth-last syllable in a word. Similarly, there are no syntactic rules which require movement of a constituent over, say, three adjacent constituents (Nessel 2006a: 1375). If grammars are unable to count then this poses a problem for Steinmetz’s theory, in which counting is a necessary part of the principle of gender tally. Rice (2006: 1408), however, illustrates that counting can be avoided in OGAT.

Prince & Smolensky (2004: 32, 75f., 258) assert that counting is not necessary in OT, since sets of violations can simply be compared instead of counted.¹⁴ Faced with, for example, two candidates, in order to determine which has the fewest violations, i.e. is optimal, the two sets of cells are compared firstly to see if either of them is empty. With OGAT, this applies to the whole GENDER FEATURES block due to the crucial non-ranking of these constraints, but in an OT scenario where all constraints are individually ranked, this applies only to the first cell. If one of the cells is empty then this candidate is optimal. If neither of them are empty, however, then one violation is removed or ‘cancelled’ from each candidate and the process starts anew. This cancellation process then continues to be applied until an empty cell is found. Should both candidates’ cells be empty, then the next cells along must be

¹⁴ See Prince & Smolensky (2004: 258) for a detailed proof of this claim.

compared. In the case of OGAT, this involves the default markedness hierarchy, to which the cancellation process also applies. Importantly, the cancellation process means that the optimal candidate can be identified without referring to the total number of violations each candidate incurs, i.e. without counting.

Technically this method of cancellation, or non-counting, could also be applied to Steinmetz’s theory, but Steinmetz does specifically propose counting as a fundamental part of his theory, namely in gender tally, which he expresses as the instruction: ‘count the number of times each gender is assigned and assign the noun the gender with the highest value’ (Steinmetz 1986: 193). Rice’s (2006) model therefore seems advantageous for this reason.

1.4.5.2 Strengths of OGAT

The main advantage of OGAT over the aforementioned approaches to principle interaction (namely semantics » form; form » semantics; morphology » semantics » phonology; and individual rule ranking) is that it is able to account for many of the examples which prove problematic for the other approaches. This leads Rice (2004: 1) to claim that OGAT is ‘the only approach which can successfully analyse the gender assignment conflicts arising in an analysis of German’.

For instance, OGAT is able to correctly predict the genders of *Pflanze*, *Waffe*, *Frucht*, *Gymnastik*, *Wacholder* and *Mohn*, which are problematic for the semantics » form approach; *Bote*, *Junge*, *Euro*, *Mango*, *Tequila*, *Wodka*, which are problematic for the form » semantics approach; *Pflanze*, *Waffe*, *Frucht*, *Wacholder* and *Mohn*, which are problematic for the morphology » semantics » phonology approach; and *Gemüse*, *Gemeinde*, *Pflanze* and *Waffe* which are problematic for the individual rule ranking approach. The OGAT tableaux for *Mohn* and *Mango*, for example, are shown in (9).

(9)

	GENDER FEATURES			DEFAULT HIERARCHY		
	FLOWERS→*M,*N	CVC→*F	V _{LONG} C→*F	*N	*F	*M
<i>Mohn</i> ‘poppy’						
☞ <i>der Mohn</i>	*					*
<i>die Mohn</i>		*	*!		*	
<i>das Mohn</i>	*			*!		

	GENDER FEATURES		DEFAULT HIERARCHY		
	FRUIT→*M,*N	-O→*M,*F	*N	*F	*M
<i>Mango</i>					
<i>der Mango</i>	*	*!			*
☞ <i>die Mango</i>		*		*	
<i>das Mango</i>	*		*!		

A second advantage of OGAT over the approaches in which principles are ranked according to type is that it can account for conflicts between gender assignment principles of the same type, e.g. a conflict between multiple phonological principles, which, as stated above, approaches like the semantics » form, form » semantics and morphology » semantics » phonology models that have been proposed generally cannot. For instance, OGAT is able to correctly predict the genders of both *Gemeinschaft* and *Schwur*, which, as stated in 1.4.1, are problematic for all three of the aforementioned approaches. The OGAT tableau for *Schwur*, for example, is shown in (10).

(10)

	GENDER FEATURES			DEFAULT HIERARCHY		
	$V_{LONG+HIGH} /r/$ MONO → *M, *N	CCVC → *F, *N	[[]C- MONO → *F, *N	*N	*F	*M
<i>Schwur</i> 'vow'						
☞ <i>der Schwur</i>	*					*
<i>die Schwur</i>		*	*!		*	
<i>das Schwur</i>	*	*	*!	*		

A further advantage of OGAT over the morphology » semantics » phonology approach is that it does not require the potentially difficult separation of phonological and morphological gender assignment principles (see Corbett 1991: 33), since, in OGAT, all GENDER FEATURES constraints are equal and their type is thus of no importance.

Additionally, unlike the aforementioned proposals, OGAT specifies a strategy for gender assignment in cases where a noun does not have any features which are relevant to gender assignment, namely the assignment of masculine via the default markedness hierarchy. In this sense, OGAT is a more comprehensive theory of gender assignment than the aforementioned proposals.

A further strength of OGAT is its support from language typology, a point which is highlighted by Rice (2004: 1, 2006: 1409), who argues that theories of gender assignment should offer a 'cross-linguistically robust theoretical perspective'. It is a well-established feature of OT that constraints are universal and that languages differ only in their rankings of these constraints (Prince & Smolensky 1993/2004). In terms of OGAT, this means that the gender assignment systems of all languages are hypothesised to consist of a set of GENDER FEATURES constraints (although the specific GENDER FEATURES (sub)constraints are of course language specific) and default markedness constraints, but that there will be variation in the rankings of these in different languages.

Proponents of OT state that the set of all possible constraint rankings yields a typology of all possible grammars (a "factorial typology"). This means that the grammar of every known human language must correspond to one of these possible rankings (it does not, however, necessarily mean that every possible ranking must be instantiated by an observed human language (e.g. Kager 1999: 35), as Rice

(2006: 1409) believes). Therefore, evidence of languages whose gender assignment systems have one of the possible rankings would strengthen the credibility of the factorial typology given by the constraints of OGAT. Such evidence is provided by Rice (2006: 1409).

Firstly, Rice (2006) claims that there are languages in which each of the three genders found in German has been argued to be the least marked, i.e. default. He maintains that masculine is the least marked in German, Norwegian, Russian and French. Neuter is claimed to be the default in Icelandic (Steinmetz 1985, 2006) and feminine is proposed as the least marked gender in Kala Lagaw Ya (Corbett 1991: 11). Although it might be argued that one cannot equate the category “masculine” in German with “masculine” in Russian, for instance, since the names of gender categories are in a sense arbitrary, the comparisons drawn here are based upon the fact that “masculine” is the category to which nouns denoting men and male animals typically belong, and “feminine” is the category to which nouns denoting women and female animals typically belong.

An additional possible ranking in the factorial typology of OGAT is one in which GENDER FEATURES constraints appear amongst the default markedness constraints. Rice (2006: 1409) claims that there are also languages in which this occurs. For example, in a language with no neuter nouns, e.g. French, Rice (2006: 1409) assumes that the constraint *NEUTER would be ranked above the GENDER FEATURES constraints, yielding a ranking such as *NEUTER » GENDER FEATURES » *FEMININE » *MASCULINE. This would mean that a fatal violation is immediately assigned to every neuter candidate. Such typological support adds plausibility to OGAT.

A final strength of OGAT to be mentioned is Nessel’s (2005, 2006a) claim that OGAT is compatible with general cognitive and linguistic principles. For example, Nessel (2005: 172, 2006a: 1376f.) states that OGAT, specifically its “gender tally” mechanism, receives support from connectionist processing (e.g. McClelland & Elman 1986). In a connectionist model, when a target (in this case, a noun) activates particular units (in this case, assignment rules) in a network, one of the factors determining the selection of a unit is the degree of conceptual overlap. The more overlap, the more likely a unit is to be selected. Nessel explains that this is parallel to the gender tally mechanism in OGAT, since in cases where a majority of principles compete with a minority, the majority represents the higher degree of overlap.

Furthermore, Nessel (2005: 173ff., 2006a: 1379f.) highlights that the mechanisms of OGAT are compatible with Kiparsky’s (1982) Elsewhere Condition. The Elsewhere Condition (also known as the Paninian Principle and the Subset Principle) states that the application of a specific rule overrides the application of a more general rule. This is consistent with OGAT, since in OGAT, specific GENDER

FEATURES constraints are ranked above the constraint which causes the default gender to be assigned (masculine in German).

In summary, OGAT appears to be a promising approach to German gender assignment which merits further investigation. This is because the theory seems better able to account for the German data than the aforementioned approaches to principle interaction. The theory also does not require the potentially problematic splitting of morphological and phonological principles. Additionally, OGAT provides an explicit strategy for gender assignment when there is a conflict between principles of the same type and in cases where there are no applicable principles at all. Furthermore, OGAT is claimed to receive support from language typology (Rice 2006) and from general linguistic and cognitive principles (Nesset 2005, 2006a). There are, however, several aspects of OGAT that require further exploration. These will be outlined in the next section.

1.4.5.3 Weaknesses and aspects of OGAT for further investigation

The first weakness and aspect of OGAT which requires further investigation concerns the amount of data upon which the proposal is based. Both Rice and Steinmetz's theories are based on relatively narrow data sets, meaning that neither provides a complete analysis of German gender assignment. Steinmetz (1985, 1986, 2001, 2006) proposes around 30 gender assignment rules in total which are used to account for only a few categories of nouns, such as names of fruits, items of clothing and musical instruments. Rice's (2006) OGAT proposal for German involves just three GENDER FEATURES constraints which are argued to account for the genders of just a handful of nouns. Consequently, further investigation of the theory is required in order to determine whether it is able to successfully predict the genders of a significantly larger sample of German nouns, as will be done in this study.

The second weakness and aspect of OGAT for further investigation is the constraint set. As demonstrated above, the outcome of an OGAT analysis is entirely dependent on the composition of the constraint set, and therefore all claims made above concerning the theory's predictive powers are wholly reliant on the proposed constraints being valid. However, OGAT does not specify what constitutes a valid GENDER FEATURES constraint. This criticism also applies to the other models of principle interaction discussed, since they also do not stipulate any requirements of a valid gender assignment principle. This means that, in theory, any constraint can be proposed ad hoc in order to ensure that an OGAT analysis yields the correct result, a point alluded to by van Berkum (1996: 40).

Indeed, a number of dubious constraints have been proposed as part of OGAT, arguably for this very reason, i.e. as 'postfactum rationalisations' (Comrie 1999: 461). For instance, Steinmetz (2006: 1436)

proposes the constraint F.HOLL. → N ‘a noun which designates a functional hollow is neuter’. A functional hollow is defined by Steinmetz (2006: 1434) as ‘a disk or [...] a complete or partial enclosure, whereby the hollow portions thereof are functional in that they are criteria for defining the object in question’, for example: *das Rad* ‘wheel’, *das Joch* ‘yoke’, *das Gesicht* ‘face’, *das Ohr* ‘ear’ and *das Ei* ‘egg’. Steinmetz (2006: 1434f.) admits that this is a ‘disquietingly subtle’ notion, yet argues that there is strong typological motivation to propose such a constraint, in that the aforementioned examples of functional hollows are all neuter across many, if not all, Germanic and Slavic languages. He states that the alternative to including this constraint would be to simply assume that this case of cross-linguistic assignment of neuter is incidental.

However, inclusion of this constraint seems questionable for a number of reasons, primarily due to its conceptual ambiguity and also because there appear to be quite a number of counterexamples, such as *der Mund* ‘mouth’, *der Magen* ‘stomach’, *der Ring*, *der Kranz* ‘wreath’, *der Sack* ‘sack/bag’, *der Brunnen* ‘well’ and *der Bagel*. Of course, the masculine gender of these counterexamples could be a consequence of the interactions of other constraints, but there do nevertheless seem to be numerous functional hollows which are not neuter. As a result, the link between functional hollows and the neuter gender seems tenuous. The key problem, however, is that because there are no established limits on OGAT constraints, constraints like this one can be easily included in an analysis.

The obscurity surrounding the question of what constitutes a valid constraint or indeed any linguistic rule or principle has been recognised by Enger (2009: 1290), who claims that the issue has been ‘shunned’ by those working on gender assignment and also avoided by linguists in other disciplines, as noted for instance by Lass (1984: 227) for rule-based phonology. In terms of standard OT, McCarthy (2007: 8f, 2008: 166) also acknowledges that OT itself offers very little guidance about the composition of constraints. Enger (2009) highlights that this problem is particularly relevant to the many semantic GENDER FEATURES constraints that have been proposed (seemingly because these tend to be less straightforward than the majority of formal constraints), and especially to those semantic constraints which do not refer to the “semantic core”, i.e. nouns with animate referents whose gender is assigned according to the NGP (see 1.3.3.1).

Enger (2009) labels such non-core constraints ‘crazy rules’,¹⁵ a term coined by Bach & Harms (1972) to describe attested rules for which there is no apparent synchronic motivation. It was originally used in the context of phonology for rules which do not appear to have any phonetic grounding, often because

¹⁵ Enger (2009) uses the term ‘rule’ throughout his paper, even when discussing OGAT, despite the fact that OT has dispensed with rules and deals with constraints only.

the original phonetic conditioning factor(s) have since disappeared, making such rules appear ‘crazy’. Enger (2009: 1290) thus uses the term for any GENDER FEATURES constraint which ‘lack[s] grounding in the semantic core of gender’, which also includes all morphological and phonological constraints.

Enger (2009) by no means dismisses all “crazy” GENDER FEATURES constraints, but recognises that some are more plausible than others, and suggests that a set of guidelines needs to be established in order to determine which of the proposed “crazy” constraints are plausible and should consequently be included in an OGAT analysis. Accordingly, Enger (2009: 1290-1294) proposes six criteria by which constraints should be assessed in order to determine their validity:

- i. A plausible rule should cover a reasonable share of the possible candidate nouns.
- ii. If a rule has parallels in other languages, that is an advantage.
- iii. If the rule invokes a particular “semantic feature”, it is an advantage if we recognise this feature from other parts of the grammar of the language in question.
- iv. If the rule invokes a particular “semantic feature”, it is an advantage if we recognise this feature from elsewhere, and it should be reasonably clear what nouns the rule covers.
- v. If rule A is necessary in order to account for exceptions from a well-established rule B, this is indirect evidence for A.
- vi. If a rule is diachronically productive (i.e. if new nouns are assigned the gender predicted by a particular “crazy rule”, or if old nouns change gender in the direction that a particular crazy rule predicts), this is an indication that this rule is valid.

These criteria proposed by Enger (2009) seem promising, since their application provides a practical method for determining whether there is independent evidence for a constraint, rather than simply having to rely on the fact that its inclusion in an analysis yields the correct result, which, as we have seen in the case of “functional hollows”, does not seem sufficient. Indeed, “functional hollows” appears, at least on first glance, to fail the majority of these criteria, which, if true, would provide us with a principled reason to reject this rather dubious constraint. Additionally, some of the criteria proposed can also be used to determine the validity of constraints relating to the form of a noun, such as the criterion demanding that a principle cover a reasonable share of possible candidate nouns.

Moreover, some of the criteria proposed by Enger (2009) coincide with some general guidelines about OT constraints outlined by McCarthy (2002: 42), such as the requirement of an OT constraint to be simple, which coincides with the latter part of Enger’s (2009) fourth criterion, the requirement of an OT constraint to account for the observations, which coincides with Enger’s (2009) first criterion, and the requirement of an OT constraint to have typological justification, which coincides with Enger’s second criterion.

However, in order to be of full practical use, specific methods for assessing constraints against each of the criteria will need to be developed, since Enger (2009) does not offer any practical suggestions on how this might be done. For instance, with his first criterion, Enger (2009) does not provide any details as to what constitutes a 'reasonable share' of candidate nouns, nor does he specify the number of exceptions a principle would have to account for in order to fulfil the fifth criterion.

To summarise, the fact that OGAT in its current state does not specify any restrictions on its constraints is a huge problem for the theory. However, this could be rectified by the application of a set of criteria which stipulate certain properties of GENDER FEATURES constraints, such as those proposed by Enger (2009). This would mean that a finite set of constraints could be established for German. Enger's (2009) suggestions of criteria seem promising as they provide a principled way of assessing dubious constraints such as "functional hollows are neuter". Nonetheless, his criteria will need developing in order to establish specific methods for their application. Enger's (2009) criteria will serve as the foundation for Chapter 2, where they will be discussed further.

A third aspect of OGAT which warrants further investigation is the proposed ranking of constraints within the default markedness hierarchy. As stated in 1.4.5.1, Rice (2006) proposes that the default hierarchy for German consists of the constraints *NEUTER » *FEMININE » *MASCULINE to reflect the fact that neuter is the most marked and masculine is the least marked gender. His main reason for claiming masculine to be the default gender is the 'preponderance of masculine nouns' in German (Rice 2006: 1406). Köpcke & Zubin (2009: 148ff.) label this notion of default gender the "stochastic default". Rice (2006) bases his claim on Köpcke's (1982) finding that approximately two thirds (64%) of the 1466 monosyllabic nouns listed in the Duden are masculine.

However, this is an interesting justification, since, not only is Köpcke's (1982) estimate based only on monosyllables, but, of the 1466 monosyllables analysed, there are more neuter nouns (22%) than feminine (14%). Therefore, if the markedness hierarchy is to truly reflect this data, it should read *FEMININE » *NEUTER » *MASCULINE. Consequently, it is important to consider other estimates of the distribution of nouns across the genders in German in order to determine whether the markedness hierarchy can indeed be justified by category size.

Source		Sample size	Masc.	Fem.	Neut.
Wängler (1963, cited in Hoberg 2004: 83)	Rangwörterbuch der hochdeutschen Umgangssprache	200	41%	37%	22%
Meier (1964), based on Kaeding (1897/1898, cited in Hoberg 2004: 83)	Deutsche Sprachstatistik: gehobenes Schriftdeutsch des ausgehenden 19. Jahrhunderts	500	35%	46%	19%
Oehler (1966, cited in Hoberg 2004: 83)	Grundwortschatz	953	39%	39%	22%
Arndt (1970: 252)	Dialogue corpora of 7 German novels from 1930-1955	ca. 300	40%	30%	30%
Bauch (1971: 417, cited in Wegener 1995: 107)	First 500 nouns of a literary work by Heiduczek	500	49%	32%	19%
Augst (1975: 37, cited in Hoberg 2004: 82)	Kernwörter	2162	67%	13%	20%
Rosengren (1977, cited in Hoberg 2004: 83)	Frequenzwörterbuch der deutschen Zeitungssprache	500	40%	44%	16%
Ruoff (1981, cited in Hoberg 2004: 83)	Häufigkeitwörterbuch gesprochener Sprache	500	46%	32%	23%
Köpcke (1982)	Monosyllables	1466	64%	14%	22%
Schiller & Caramazza (2003: 171)	Monomorphemic nouns in the CELEX database (Baayen et al. 1995)	4113	43%	38%	19%
	Token frequency in the above		39%	35%	26%
Duden Sprachwissen (2017)	Duden corpus	ca. 4.5 billion	34%	46%	20%

Figure 3: Estimates of the distribution of nouns across the three genders in German

As shown in Figure 3, there is considerable variation among the sources in their estimates of category size. This is most likely due to the varied data and sample sizes upon which the estimates have been made, as well as the fact that some of the sources calculate token rather than type frequency.

However, in spite of the variation, the majority of the sources (8/11) do report masculine to be the largest or joint largest category and neuter to be the smallest category. The greatest variation seems to occur in the estimates of feminine, with only around half of the sources reporting it to be the second largest category, two reporting it to be the smallest, and four as the largest or joint largest. It thus seems that, based on category size alone, the hierarchies *N»*M»*F or *F»*N»*M might also be possible.

Wegener (1995: 62) highlights that the notion of a stochastic masculine default is supported by evidence from loanwords, since many studies report that the majority are assigned masculine, such as Carstensen (1980) and Callies et al. (2012: 68), as well as most of those shown in Figure 4. See also section 3.3.1.4.

	Source	Sample size	Masc.	Fem.	Neut.	MGNs
Arndt (1970)	Nouns frequently used in US English tested on 25 native German speakers	130	35%	50%	15%	-
Yang (1990)	English loanwords in 24 issues of <i>Der Spiegel</i> (1950-1980)	1204	60%	16%	24%	-
Schulte-Beckhausen (2002)	English monosyllables in various German dictionaries (1852-1999)	402	53%	8%	19%	20%
Chan (2005)	English loans in <i>Duden: Das große Wörterbuch der deutschen Sprache</i> (1999)	3105	48%	16%	30%	6%
Onysko (2007)	English loans in <i>Der Spiegel</i> corpus (2000)	1023	61%	18%	20%	1%

Figure 4: The distribution of English loanwords across the three German genders

The only study which does not report a majority of masculine loans is Arndt (1970), for whom the hierarchy *N»*M»*F would best reflect the data. The remaining studies all report a masculine majority, and interestingly report neuter to be the second largest category, suggesting the hierarchy *F»*N»*M.

Rice (2006: 1406) also provides further justification for a masculine default in German, stating that ‘the number of principles required to cover all the nouns in this category is far greater than the number of principles needed for neuter’. Presumably Rice (2006) is arguing that since such a high proportion of gender assignment principles assign masculine, it must be the default gender. So whilst the gender of feminine and neuter nouns can be accounted for by a small subset of principles, the entirety of the elsewhere category constitutes masculine nouns. This idea of a default has been labelled the “system default” by Köpcke & Zubin (2009: 148ff.), which reflects Kiparsky’s (1982) Elsewhere Condition, given that masculine is assigned unless a more specific principle assigns feminine or neuter. Also in favour of a masculine “system default” in German are Eisenberg (2013: 490) and Wegener (1995: 62), who both claim that masculine is assigned in cases where there is no particular reason to assign another gender.

Experimental evidence from Hohlfeld (2006: 138) corroborates this claim. As part of her experiment on pseudo-nouns, a control test was undertaken in which participants assigned genders to pseudo-nouns which had no features that are known to be relevant to gender assignment.¹⁶ The results showed a bias towards masculine (50%), followed by neuter (40%) then feminine (10%), which would suggest the “system default” hierarchy *F»*N»*M.

¹⁶ Although whether this is indeed the case for every single test item is unclear, given that among them were *Fusem*, *Tapam* and *Schogem* (where final stressed *-Vm* has been argued to be associated with neuter (Nelson 1998: 215), *Gumdul* and *Bandul* (whose final syllable might be equated with the pseudosuffix *-el*), and *Gischep* (whose initial syllable might be equated with *Ge-*).

It has also, however, been claimed that the “system default” in German might instead be argued to be neuter. Köpcke & Zubin (2009: 148ff.) state that this is because neuter is assigned in the following cases, in which, they argue, no specific gender assignment principles apply: cases of discourse anaphora, e.g. *der Wagen war umgestürzt, und das hatte ihn erschüttert* ‘the car had turned over, and that shook him’; exophoric reference to an unknown object, e.g. *was ist denn das?* ‘what’s that?’; headless NPs without a specific referent: *das Schöne* ‘the beautiful thing’; presentative structures, e.g. *das ist meine Mutter* ‘this is my mother’; and expletive subjects, e.g. *es scheint* ‘it seems’. Similar arguments are also made by e.g. Talanga (1987: 93), Chan (2005: 75-79) and Fries (1997: 18f.).

However, Köpcke & Zubin (2009: 148ff.) do also remark that some of these cases are often analysed as having “genderfree” pronouns and determiners rather than ones which are grammatically neuter. Additionally, Corbett & Fraser (2000: 69f.) state that such reasoning only applies to the sentence default rather than the default gender of nouns themselves. That is to say, for German, neuter might be seen as the default on a sentential level, but masculine is the default for nouns.

Finally, it is also claimed by Köpcke & Zubin (2009: 148ff.) that feminine could be argued to be the default gender in German, namely on the basis of morphological productivity. They state that, whilst productive derivational processes lead to the formation of nouns of all three genders, feminine nouns have the highest type and token frequency of all derived nouns (Köpcke & Zubin 2009: 148ff.). Hoberg (2004: 82), too, observes that, in terms of type frequency, the majority of suffixes create feminine nouns, followed by masculine and then neuter, which would suggest the hierarchy *N»*M»*F.

In sum, it seems that a number of default hierarchies could potentially be justified for German in addition to the *N»*F»*M hierarchy proposed by Steinmetz (1985, 1986 etc.) and Rice (2006). In line with the notion of a “stochastic default”, the hierarchies *N»*M»*F or *F»*N»*M also seem plausible. In terms of a “system default”, neuter might be seen as the default gender (*F»*M»*N or *M»*F»*N), and as a “morphological default”, the hierarchy *N»*M»*F would be justifiable. It is therefore a matter for further investigation which of these hierarchies yields the highest number of correct predictions. This aspect of OGAT will be tested in 3.4.3.

A final area for further investigation are two features of Steinmetz’s (1985, 1986 etc.) rule-based theory which have not explicitly been integrated into Rice’s (2006) OGAT model, namely “strong rules” and “exceptional gender”.

“Strong rules” are defined by Steinmetz (1986: 194) as rules which ‘erase all others from the gender table’, for example ‘-ung = f’. That is to say, if a strong rule applies to a noun, then the noun will automatically receive the gender that the strong rule assigns to it, regardless of any other rules which may apply to the noun. However, this notion appears to be incompatible with OGAT because it would require some GENDER FEATURES constraints to be ranked higher than others.

While Rice (2006: 1409) states that this is a matter for future research, he argues that “strong” constraints such as $-UNG \rightarrow *M, *N$ would not necessarily be incompatible with OGAT, since there is a wider pattern of the rightmost element of complex nouns determining their gender (i.e. the LMP, see 1.3.1.4). He claims that ‘these cases presumably do not require explicit reference to language specific material, but rather refer to a morpheme’s status as a head’ (Rice 2006: 1409). It is, however, unclear how this would work without reference to language-specific material, because it is not the case that every German suffix is “strong”, e.g. the suffix *-nis* does not categorically assign neuter, as seen in 1.3.1.1.

Moreover, Rice’s (2006) claim would only account for “strong” morphological rules. Technically Steinmetz’s proposal of strong rules could also be extended to semantic or phonological constraints, which would not be accounted for by headedness. How OGAT is able to account for “strong rules” is therefore a matter for further investigation and will be discussed in 3.3.3.

The second aspect of Steinmetz’s theory which is not integrated into Rice’s (2006) proposal of OGAT is that of “exceptional gender assignment”. Steinmetz (1986: 209) asserts that, although his theory is able to account for the genders of many nouns, ‘it is by no means the case that rules now known cover all the data’. He states that it is entirely possible that the German gender assignment system may in parts be irregular, with a small number of nouns receiving a gender which is not systematically assigned. Examples of nouns which Steinmetz claims to exhibit exceptional gender assignment are *das Auge* ‘eye’, *die Post* and *das Knie* ‘knee’ (Steinmetz 1986: 209, Steinmetz & Rice 1989: 175, Steinmetz 2006: 1432). It therefore remains to be seen how many more nouns in the German lexicon have “exceptional gender”, and whether there are any commonalities among exceptional nouns. This will be discussed further in 3.5.2.

In sum, there are several areas of OGAT which require further investigation. The first concerns testing OGAT on a larger sample of nouns in order to more thoroughly ascertain its predictive powers. The second involves independent regulation of the constraint set in order to ensure the credibility of any results obtained from testing OGAT. The third consists of testing variations of the default hierarchy to determine which yields the best results, and the fourth involves investigation of Steinmetz’s (1985,

1986 etc.) claims about strong rules and exceptional gender. The next section will focus on previous investigation of OGAT in the gender assignment literature.

1.4.5.4 Previous testing of OGAT

There has been relatively little exploration of OGAT in the literature, with just a handful of studies having been completed to date. This section outlines three studies on OGAT and German – the first two by Onysko (2007) and Burkhard (2013) focussing on English loanwords and the third by Scheibl (2008), and three studies on OGAT and other languages – one on Italian by Thornton (2009), one on Catalan by El-Youssef (2010), and one on Russian by Galbreath (2010).

1.4.5.4.1 OGAT and German

Onysko (2007)

Onysko (2007) tests OGAT on 1,017 English loanwords in German, taken from a tagged portion of *Der Spiegel 2000* corpus. He claims that OGAT is able to correctly predict the genders of the loans in the corpus. Onysko (2007: 156) argues that OGAT is a ‘valid approach’ to German gender assignment because it provides a comprehensive account of gender assignment which makes it ‘methodologically superior’ to approaches involving the application of a set of disconnected rules. He also cites the capacity of OGAT to account for gender assignment across the German lexicon and in other Indo-European languages as a great strength of the theory.

Onysko claims that his results demonstrate that OGAT can account for the gender of English loanwords in German, since the genders of all loanwords in the corpus are able to be predicted either by one or more GENDER FEATURES constraints or by the default markedness hierarchy. He states that the majority of masculine loanwords in the corpus have no specific gender-assigning features and therefore their masculine gender can be explained by the default hierarchy (Onysko 2007: 163, 178).

However, a major problem with Onysko’s analysis is that he does not specify any limits on his constraint set, as is the case with the original proposals of OGAT by Steinmetz (1985, 1986, 2001, 2006) and Rice (2006). This means that a number of dubious “crazy” semantic constraints are included in Onysko’s analysis, such as those involving the allegedly feminine-assigning features: ‘production and processing sites’ and ‘semantic diminutives’, and those involving the allegedly neuter-assigning features: ‘sheets, strips and slabs’, ‘mesh-like structures’, and ‘cushions and reclining’.

Most, if not all of these constraints are unlikely to fare well by Enger’s (2009) criteria. For instance, the constraint “semantic diminutives are feminine” seems problematic for a number of reasons. This

constraint was originally proposed by Steinmetz (1986: 202f.), who uses the term 'semantic diminutive' to denote 'smaller, weaker, less active, less prominent' referents, such as *die Bö* 'gust' (as opposed to *der Wind*), *die Insel* 'island' (as opposed to *das Land* 'land, country'), *die Tür* 'door' (as opposed to *das Tor* 'gate') and *die Nadel* 'needle' (as opposed to *der Nagel* 'nail').

There appear, however, to be many semantic diminutives which are not feminine: *der Junge* 'boy' (as opposed to *der Mann* 'man'), *das Baby* 'baby' (as opposed to *der Erwachsene* 'adult'), *das Insekt* 'insect' (as opposed to *das Großtier* 'large animal'), alongside all the neuter diminutive nouns with the suffixes *-chen* or *-lein*, e.g. *Brötchen* 'bread roll', *Büchlein* 'booklet'. Moreover, the concept of a semantic diminutive is extremely ambiguous. It seems that many things could be construed as a semantic diminutive when compared to a larger object. For example, Steinmetz (1986: 202f) claims that *die Bö* 'gust' is a semantic diminutive when compared to *der Wind*, but by the same logic *der Wind* could then be argued to be a semantic diminutive if compared to *der Sturm* 'gale' or *der Tornado*. In his analysis, Onysko (2007: 170) claims that this constraint applies to *die Couch*, but it is unclear what *Couch* is a diminutive in relation to. Onysko's inclusion of such dubious constraints in his analysis therefore brings the validity of his results into question.

Burkhard (2013)

Like Onysko (2007), Burkhard (2013) also tests OGAT on English loanwords in German. However, contrary to the findings of Onysko, Burkhard claims that OGAT cannot always produce the correct gender predictions for English loanwords in German. This is probably because the methodology used by Burkhard is quite different to that of Onysko.

In Burkhard's (2013) study, 37 German speakers were asked to assign genders to 40 English loanwords. Burkhard then tested these 40 nouns using OGAT and found that the gender predictions made by OGAT did not always match the genders assigned to the 40 loans by the majority of the participants of the study. There are, however, two main problems with Burkhard's methodology.

Firstly, the 37 German speakers who participated in the study consisted of 23 native German speakers and 14 'advanced learners' of German. An 'advanced learner' is described by Burkhard (2013: 19) as someone with at least five years of university-level German instruction. The problem with this is that a non-native speaker's intuition is not generally considered to be completely reliable. The main purpose of OGAT is to model the gender assignment system of a native speaker. It therefore seems irrelevant if the prediction made by OGAT does not match the prediction made by a non-native speaker of German.

Secondly, Burkhard uses just six GENDER FEATURES constraints in her OGAT analysis – five formal constraints and one semantic constraint. It is quite likely, however, that there are in fact more constraints which are relevant to the loanwords in her sample, which would consequently change the outcome of the analysis. Burkhard (2013) also includes a lexical equivalence constraint in her analysis (stating that loans adopt the gender of their nearest lexical equivalent), which she ranks below the GENDER FEATURES block without any justification. It could therefore be the case that the ranking of this constraint is not correct or that her method of determining the nearest lexical equivalent is problematic (the potential problems with lexical equivalence have been highlighted by e.g. Onysko (2007: 166ff.) – see further discussion in 3.3.1.2).

When looking at Burkhard's results, it can be seen that OGAT is in fact able to correctly predict the genders of most of the nouns discussed in the paper. OGAT only fails to select the correct gender for three nouns: *Farm*, *Mall* and *Totebag*. At least twelve of the nouns in the sample are not discussed, however. Additionally, Burkhard's results show that the majority of simplex loanwords in her sample are assigned masculine, which supports the claim made by both Rice (2006) and Steinmetz (1985, 1986 etc.) that masculine is the default gender in German.

One of Burkhard's additional criticisms of OGAT is that it is unable to account for the variation present in her results. She highlights that, due to only being able to select one optimal candidate, OGAT cannot reflect the fact that some participants allowed more than one gender for some of the loans in the sample. However, one possible explanation which could be accounted for in OGAT is that the variation was due to participants selecting different lexical equivalents for the loans. Burkhard (2013: 44) herself comments on the variants *der Cupcake* and *das Cupcake*, observing that the participants who selected *der Napfkuchen* as its lexical equivalent assigned masculine to the loan and those who selected *das Törtchen* assigned neuter. These two variants could therefore be accounted for by the differing violations of the lexical equivalence constraint. This idea is discussed further in 3.3.1.2 and 3.3.4.8.

Scheibl (2008)

Scheibl (2008: 67f.) tests OGAT on a small selection of German nouns and subsequently rejects the theory based on the claim that it is unable to account for the genders of hybrid or proper nouns.

In terms of hybrid nouns, Scheibl highlights three cases which he argues are problematic for OGAT, namely: *Exzellenz*, *Majestät* and *Mädchen*.

It is assumed that Scheibl considers *Exzellenz* 'excellency' and *Majestät* 'majesty' problematic for OGAT in cases where they refer to men, causing a conflict between their (contextual) masculine

meaning and their feminine form. In OGAT, this should lead to the assignment of masculine gender, yet these two nouns are feminine.

However, it is important to consider that *Exzellenz* and *Majestät* as nouns do not inherently refer to people; they simply denote the quality of being excellent or majestic. They are used to refer to people only when preceded by a possessive pronoun, which converts them into honorific titles. *Seine Majestät* is simply a substitution for the name of the referent. When titles or epithets are used to replace names, it is known as *antonomasia*. We see a similar case with *der stählerne Schmetterling* ‘The Steel Butterfly’ (the nickname given to Imelda Marcos), in which *Schmetterling* itself does not refer to a person and therefore, the societal gender of its referent when used *antonomastically* does not play a role in the assignment of grammatical gender to *Schmetterling*. In the same way, the feminine gender of *Majestät* and *Exzellenz* is decided on the nouns’ inherent features (namely the presence of the suffixes *-tät* and *-enz*) alone. The fact that these nouns can be used *antonomastically* to refer to men does not play a role in the determination of their gender, which nullifies Scheibl’s argument.

As for *Mädchen* ‘girl’, Scheibl (2008: 67) claims that if the noun is subject to the two principles “nouns in *-chen* are neuter” and “nouns denoting women are feminine”, then OGAT would incorrectly predict feminine for *Mädchen*. There are, however, some potential adjustments to this constraint set which would resolve this issue.

The first would be to specify that “nouns denoting women are feminine” refers indeed to only women, i.e. adults, since, as noted in 1.3.3.1, young women do not seem to be assigned gender according to the NGP in German (e.g. Zubin & Köpcke 1981: 445, Köpcke & Zubin 1996: 483). The second would be to add a constraint to the constraint set, namely “nouns denoting young humans (and young animals) are neuter”, which, as suggested in 1.3.3.1, seems to form part of the NGP in German (e.g. Zubin & Köpcke 1981: 444f., Flämig 1991: 456, Eisenberg 1999: 156, Hoberg 2004: 102, Weinrich 2007: 336). With either of these adjustments to the analysis, OGAT would be able to predict the correct gender for *Mädchen*. A further option would be to explain the gender of *Mädchen* by claiming that *-chen* is a “strong rule” (see 1.4.5.3 and 3.3.3).

In terms of proper nouns, Scheibl (2008: 68) lists three cases which he deems problematic for OGAT: *das (schöne) Salzburg* (city), *das Dreher* (beer) and *die Prinz Charles* (ship). He claims that, due to the default markedness hierarchy, the feminine form of *Salzburg* (namely the element *Burg*) should outrank its neuter meaning as a city, predicting **die (schöne) Salzburg*. Similarly, the masculine form of *Dreher* (namely its final *-er*) should outrank its neuter meaning as a beer, predicting **der Dreher*, and the masculine properties of *Prinz Charles* should outrank its feminine meaning as a ship.

However, while on the surface these appear to be valid counterexamples, Scheibl does not consider that proper nouns as inputs to OGAT may not be analysable by native speakers in the same way as common nouns.

To deal firstly with *Salzburg*, for example, Hoberg (2004: 106f.) claims that city names are neuter regardless of their form because they are analysed as wholes. This is because *Salzburg* is not a kind of *Burg*_{FEM} 'castle', in the same way that *Michelbach* does not denote a *Bach*_{MASC} 'stream' and *Friedrichshafen* is not a kind of *Hafen*_{MASC} 'port'. It is therefore unlikely that there would be a formal constraint assigning feminine to *Salzburg* due to the feminine gender of *Burg*, since it is not an analysable compound in present-day German. The only constraint that is relevant to *Salzburg* is thus CITIES→*M,*F, which would lead OGAT to yield the correct prediction.

As for *die Prinz Charles* and *das Dreher*, it might also be argued that the formal properties of these proper nouns do not play a role in gender assignment, since, like cities, they are unanalysable names. This would explain the assignment of feminine to ships with the names: *die Polarstern* (despite *der Stern*), *die Prinz Hamlet*, *die General San Martin* and *die Kaiser Wilhelm der Große* (despite being men's names) and *die Bremen*, *die Berlin* and *die Leipzig* (despite being city names). It would also explain the neuter gender of numerous varieties of beer: *das Hefeweizen* (despite *der Weizen*), *das Corona* (despite final *-a*), *das Carlsberg* (despite *der Berg*) and *das Export* (despite *der Export*). This point will be discussed further in sections 3.3.4.1 and 3.3.4.2.

1.4.5.4.2 OGAT and other languages

Thornton (2009)

Thornton (2009) tests OGAT on a small selection of nouns in Italian and ultimately concludes that the theory cannot account for Italian data. Thornton firstly tests the theory on six relatively recent masculine loanwords. The nouns all denote men, which is associated with masculine gender in Italian, and they also all end in /a/, which is associated with feminine. OGAT is able to correctly predict that these nouns are masculine.

Thornton then tests OGAT on three feminine nouns which denote women (associated with feminine) and also end in /o/ (associated with masculine). OGAT does not make the correct prediction for these nouns. However, Thornton (2009: 22) herself concedes that all three nouns are problematic examples. The first, *soprano*, is in fact used significantly more often with masculine targets than feminine targets (Thornton 2009: 22). The second, *squillo* 'call girl', is actually the elliptical version of *ragazza squillo*, where the head noun is the feminine *ragazza* 'girl' (Thornton 2009: 22f.). The third example, *virago*

'man-like woman', is a Latin loanword which was feminine in Latin and did not change to masculine like many other Latin loans ending in /o/, possibly because of its semantics (Thornton 2009: 23).

Before rejecting OGAT, Thornton tests whether the theory can make the correct gender predictions with a small adjustment, namely deleting all masculine-assigning constraints. Thornton's reasoning behind this is that masculine is the default gender in Italian and so these rules could be argued to be redundant. With this adjustment, OGAT is able to correctly predict the feminine gender of *virago*, but then cannot account for the masculine gender of the six aforementioned nouns denoting men and ending in /a/.

Thornton therefore concludes that OGAT is not a viable theory of gender assignment for Italian. Instead, Thornton advocates the ranking of semantic constraints above formal constraints, which she claims is better able to account for the Italian data. Thornton does, however, dismiss OGAT based on very little data, and partly on data which she herself considers problematic. Her conclusion is based on a total of nine nouns using just four GENDER FEATURES constraints. It is entirely possible that, firstly, there are other constraints that are relevant to the nouns tested which could consequently change the outcome of her analysis, and, secondly, that the outcome would look different if the theory were tested on a larger sample of more prototypical Italian nouns.

El-Yousseph (2010)

In El-Yousseph's (2010) investigation on gender in Catalan, she compares the predictive power of OGAT with that of an analysis in which semantic constraints are ranked above formal constraints (as advocated by Thornton (2009) for Italian, and in line with Aksenov (1984), Corbett (1991), Corbett & Fraser (2000) etc, see section 1.4.1).

El-Yousseph firstly tests the two approaches on the names of four makes of car (associated with masculine) ending in *-a* (associated with feminine). Both approaches are able to correctly predict that these proper nouns are masculine. When testing both theories on names of motorbikes, however, El-Yousseph claims that OGAT is not always able to make the correct predictions. For motorbike names which contain a masculine-assigning formal feature in Catalan, e.g. *Suzuki*, OGAT incorrectly predicts that they are masculine, whereas the approach in which semantics dominates form correctly assigns feminine to these proper nouns.

El-Yousseph therefore concludes that OGAT cannot correctly predict the genders of all nouns in Catalan, but claims that a theory in which all semantic constraints dominate all formal constraints is better able to account for Catalan gender assignment. Like Thornton (2009), El-Yousseph rejects OGAT

based on relatively little data; she tests the theory on two small groups of nouns (cars and motorbikes) using just four GENDER FEATURES constraints. El-Yousseph also does not consider the possibility that the form of certain proper nouns may not play a role in gender assignment (see above discussion of Scheibl (2008) and Sections 3.3.4.1 and 3.3.4.2).

Galbreath (2010)

Galbreath investigates gender assignment in Russian using OGAT, stating that OT is well suited for this purpose, since 'it is predicated on competition among constraints, which is the very process at work in gender assignment' (Galbreath 2010: 204). Based on his analysis, he concludes that gender assignment in Russian involves a combination of semantic, phonological and morphological GENDER FEATURES constraints together with a default markedness hierarchy. Galbreath states that the crucial equal ranking of GENDER FEATURES constraints is difficult to disprove in light of the Russian data, but nevertheless argues in favour of an adaptation to OGAT so that the GENDER FEATURES constraints are ranked such that the semantic constraints are above the morphological constraints, which are in turn ranked above the phonological constraints, with some additional hierarchical ranking among individual semantic constraints. He consequently endorses OGAT for Russian with the addition of these constraint rankings.

However, one aspect of Galbreath's analysis which is highly problematic is his lack of independent regulation of the constraint set. This leads him to include dubious constraints in his analysis such as "functional hollows are not masculine or feminine", thus bringing the validity of his results into question.

In summary, it can be seen that OGAT is a relatively underexplored theory of gender assignment. The studies that have been undertaken have mostly used very small data sets (particularly Scheibl 2008, Thornton 2009, El-Yousseph 2010 and Burkhard 2013), and have tested only specific subsets of the lexicon such as loanwords (Onysko 2007 and Burkhard 2013), hybrid nouns (Scheibl 2008) and proper nouns (Scheibl 2008 and El-Yousseph 2010). There are also a number of methodological problems with the studies, such as Burkhard's (2013) use of non-native speakers to determine native-speaker intuition. Most crucially of all, the usage of constraints in all of the studies is problematic, since there is no provision of independent justification for the constraint sets used, which has led to the inclusion of some questionable constraints in a number of the studies (particularly Onysko 2007 and Galbreath 2010).

1.5 Summary and plan for investigation

In this chapter, it has been demonstrated that gender assignment in German is a seemingly complex matter, which has led to the view that it is largely arbitrary and that the genders of individual nouns are therefore stored in the lexicon (e.g. Maratsos 1979, Pfau 2009) (see 1.2). However, various studies have revealed numerous patterns of correlation between the gender of a noun and its semantic, morphological and phonological properties, thus enabling the possibility of computation based on a “gender assignment system” in the lexicon (e.g. Altmann & Raettig 1973, Zubin & Köpcke 1981, 1986, Corbett 1991) (see 1.2 and 1.3). There is experimental evidence which supports the exploitation of a gender assignment system by native German speakers (e.g. Schiller et al. 2003, Schwichtenberg & Schiller 2004), although exactly to what extent this might occur remains a matter for further psycholinguistic investigation (see 1.2 and 1.3).

The proposed composition of the German gender assignment system is still very much a matter of debate. While it is widely agreed that it consists of semantic and formal principles, their interaction is disputed. A number of proposals for principle interaction were outlined, e.g. Spitz (1965), Köpcke (1982), Wegener (1995), but it was shown that none of these is able to account for all of the German data (see 1.4). However, a promising approach comes in the form of OGAT (Rice 2006), which claims that German gender assignment can be accounted for with a combination of crucially equally ranked GENDER FEATURES constraints and a default markedness hierarchy. OGAT seems to be able to account for many of the cases the other approaches cannot, and overcomes many of the problems associated with the other approaches (see 1.4).

There has, however, been relatively little exploration of OGAT in the literature (see 1.4.5.4), and the theory thus merits further testing. Particular aspects for further investigation include: testing OGAT using a larger sample, independent regulation of the constraint set, exploration of alternative default hierarchies, Steinmetz’s notions of “strong rules” and “exceptional gender”, and the ability of OGAT to account for the genders of loanwords, proper nouns and MGNs (see 1.4.5.3 and 1.4.5.4).

In light of the above findings, the main aim of this project is to test the predictive powers of OGAT for German by determining whether the theory is able to account for the genders of a substantial sample of nouns (introduced in 2.2). A key feature of the investigation will be to ensure that any constraints used in the study have independent justification in order for the results obtained to be valid. This will be done by assessing all relevant constraints that have been proposed using Enger’s (2009) six criteria, which will be the focus of Chapter 2. It will then be seen whether, using all independently-justified constraints, OGAT can account for the genders of the nouns in the sample, which will be the focus of Chapter 3. Chapter 3 will also investigate the ability of OGAT to account for the genders of loanwords,

proper nouns and MGNs, as well as Steinmetz's notions of "strong rules" and "exceptional gender", alternative default hierarchies and other issues arising from the sample and constraint set. Any necessary adjustments to OGAT will be made on the basis of these findings. Chapter 4 will then focus on testing the adjusted model in an experiment with native speakers using pseudo-nouns, in order to test the plausibility and psychological reality of the system. Final conclusions will be drawn in Chapter 5.

The framework to be used throughout the investigation will be OT, since this is the framework in which OGAT was proposed. However, it is important to clarify that the framework is by no means the primary focus of the study, rather it is the mechanisms of equal gender assignment principles coupled with a default hierarchy which are being tested. It is recognised that these mechanisms could equally be explored via a rule-based framework. The thesis should therefore be seen as an exploration of OGAT as a novel approach to gender assignment in German, rather than as an endorsement of the wider use of OT in linguistics.

The present study serves not only as a contribution to the gender assignment literature, but may also have wider benefits. Corbett (1991: 8) highlights that 'given the difficulty experienced by foreign learners of many gender languages, an understanding of gender assignment systems is of considerable practical importance'. Indeed, many scholars recognise the severe problems gender causes for even advanced German learners (e.g. Wegera 1997, Lemhöfer et al. 2008, Bobb et al. 2015). Since it has been demonstrated that explicit teaching of gender assignment regularities can reduce gender errors among L2 learners (e.g. Schirrmeyer 2013, Kraiss 2014), it is hoped that any findings from this investigation may benefit this group.

Chapter 2: Evaluating GENDER FEATURES constraints

2.1 Introduction

In Chapter 1, it was established that the computation of German gender by means of a gender assignment system in the lexicon is plausible given the numerous patterns of correlation between the gender of nouns and their meaning and form. However, the exact composition of this system remains a matter for further investigation. Based on initial analysis, OGAT appears a promising theory of gender assignment in German, since it seems better able to account for German data than other proposed models. Therefore, the principal aim of this project is to assess the predictive powers of OGAT for German by testing its ability to account for the genders of a sample of nouns (introduced in 2.2), which will be the focus of Chapter 3.

In order for the results of this test to be valid, however, it is first important to establish whether a set of GENDER FEATURES constraints which have independent justification can be assembled for use in the study. This is to avoid the problem outlined in 1.4.5.3 of constraints which are merely ‘postfactum rationalisations’ (Comrie 1999: 461) being included in the analysis.

Accordingly, in this chapter, all gender assignment principles that have been proposed in the literature for German which are relevant to the sample (see 2.2) will be assessed according to the six criteria proposed by Enger (2009). These criteria seek to determine whether or not there is independent evidence for a gender assignment principle, as outlined in 1.4.5.3. They are as follows:

- i. A plausible rule should cover a reasonable share of the possible candidate nouns.
- ii. If a rule has parallels in other languages, that is an advantage.
- iii. If the rule invokes a particular “semantic feature”, it is an advantage if we recognise this feature from other parts of the grammar of the language in question.
- iv. If the rule invokes a particular “semantic feature”, it is an advantage if we recognise this feature from elsewhere, and it should be reasonably clear what nouns the rule covers.
- v. If rule A is necessary in order to account for exceptions from a well-established rule B, this is indirect evidence for A.
- vi. If a rule is diachronically productive (i.e. if new nouns are assigned the gender predicted by a particular “crazy rule”, or if old nouns change gender in the direction that a particular crazy rule predicts), this is an indication that this rule is valid.

Each of these criteria will be discussed and applied to the proposed constraints in turn. As stated in 1.4.5.3, some of Enger’s (2009) criteria are vague in places, and specific methods are not suggested. Therefore, one of the aims of this chapter is to refine the criteria and develop possible methods in order to make the criteria into a testable and robust set.

At the end of this chapter, a set of constraints that emerge as valid according to the criteria will be produced, which will subsequently be used in the testing of OGAT in Chapter 3. This chapter will explore how constraints may fulfil the criteria sufficiently and those which do not will be discarded from the constraint set to be used in the analysis.

The structure of the chapter is as follows: Section 2.2 will introduce the noun sample and the constraints with which OGAT will be tested; in Sections 2.3-2.8, the constraints will be assessed according to the six criteria; in 2.9, some further issues relating to the constraint set will be discussed; and in 2.10, the final list of constraints will be presented.

2.2 The noun sample and the constraints

The sample of nouns with which OGAT will be tested consists of 592 nouns in total. These were collected by selecting the first noun (broadly defined as a word that is capitalised in Standard German) on alternate pages of the Duden *Rechtschreibung* (2013). The noun sample can be divided into two groups: the main sample consisting of 537 nouns, and a set of 55 nouns which can be deemed in some way exceptional.

The set of 55 exceptional nouns consists of: 26 nouns which are listed without an article in the Duden (22 proper nouns – 11 names of people and 11 names of places – and 4 nouns which occur only in the plural), 12 nouns which are not expected to be part of the competence of all native speakers (9 terms designated as regional and 3 archaic terms listed as obsolescent or obsolete), and 17 nouns to which multiple genders are relevant (9 multiple gender nouns, 5 homonyms and 3 adjectival nouns).¹

Where one of these 55 exceptional nouns was encountered during the sample collection, a replacement was found by selecting the next noun listed on the page, so that the main sample of 537 nouns still contains one noun from each alternate page of the Duden. In the analysis in Chapter 3, this set of 55 exceptional nouns will be dealt with separately (Section 3.3.4) to ensure that the results are not skewed in any way.

The table below shows the distribution of nouns across the genders in the main sample of 537 nouns, the sample of 592 nouns (which includes the exceptional set of 55) and the Duden corpus (Duden Sprachwissen 2017).

¹ Nouns which are recognisably loanwords were not included in the set of 55 as might be expected. This is because the sample was taken from the Duden *Rechtschreibung*, and therefore contains only loans that are established in German (which generally have an established gender), so would not be expected to significantly distort the results. Loans which have more than one possible gender, i.e. MGNs, are included in the set of 55.

	Sample of 537	Sample of 592	Duden corpus ²
Masc.	36.9%	35.1%	34%
Fem.	44.3%	40.9%	46%
Neut.	18.8%	17.6%	20%
2 genders	-	1.5%	<1.3%
masc./neut.	-	1.2%	0.9%
fem./neut.	-	0.3%	0.2%
masc./fem.	-	0.0%	0.2%
3 genders	-	0.0%	0.02%
Adj. nouns	-	0.5%	Unspecified
No article	-	4.4%	<0.1%

Overall, the distribution of nouns across the gender categories in both samples is roughly equivalent to the distribution found in the Duden corpus, indicating that they are good representations of the German lexicon. In terms of the proportion of masculine, feminine and neuter nouns, both the sample of 537 and the sample of 592 seem to reflect the proportions found in the Duden corpus relatively closely, the only exception being the slightly lower proportion of feminine nouns in the sample of 592. The proportion of multiple gender nouns in the sample of 592 is also almost the same as the proportion found in the Duden corpus. The percentage of nouns listed without an article is somewhat higher in the sample of 592 than in the Duden corpus, presumably because many proper nouns, e.g. *Berlin*, although listed without an article in the Duden, are not explicitly listed as being ‘ohne Artikel [without article]’, as others are, e.g. *Google* and *Delete*.

This full sample is listed in Appendix A.

The GENDER FEATURES constraints

The GENDER FEATURES constraints which will be assessed using Enger’s (2009) six criteria in this chapter before being used to test OGAT in Chapter 3 (if emerging as valid according to the criteria) are those which have been proposed in the literature on German gender assignment. These – many of which were outlined in 1.3 – have been either proposed originally as constraints or as gender assignment rules which have subsequently been transformed into constraints for this study.

According to McCarthy (2008: 175), ‘an OT constraint has just one job: to assign some number of violation marks to a candidate based on its output structure or how it differs from the input’. For OGAT, this means that a constraint ought to assign violation marks to candidates based on their output

² It should be noted that the statistics given by Duden about their corpus are not entirely consistent with one another, in that the percentage total amounts to 101.42% and the figures are all given to different numbers of decimal places. However, it is assumed that the percentages given for the proportion of masculine, feminine and neuter nouns (rows 1-3) have been rounded up to the nearest integer.

gender. An example of a gender assignment rule which has been converted into a constraint for this study is ‘nouns with the suffix *-ung* are feminine’ (e.g. Flämig 1991: 453, Duden Grammatik 2009: 165, Eisenberg 2013: 133). This has been transformed into the constraint $-UNG \rightarrow *M, *N$ “nouns with the suffix *-ung* cannot be masculine or neuter”, so that all candidates with suffix *-ung* that are masculine or neuter will incur a violation. This is in line with the formulation of GENDER FEATURES constraints in Rice (2006).

There are, however, some cases which are not so straightforward. It is stated by OT theorists, e.g. McCarthy (2002: 40, 2008: 175) and Nessel (2006b: 327), that constraints which ‘reproduce the effects of constraint interaction’ should be avoided (McCarthy 2008: 175). Any constraints whose definitions require “if” or “except when” clauses, for instance, are concealing ranking arguments and should be split up into two constraints (McCarthy 2002: 40, 2008: 175; Nessel 2006b: 327). This is especially problematic for OGAT given that all GENDER FEATURES constraints are to be equally ranked.

However, there are a number of gender assignment rules which have been proposed in the literature which, if simply converted into a single constraint, would incur this very problem. For example, the rule “nouns denoting non-German rivers are masculine, except those ending in schwa, which are feminine” (e.g. Flämig 1991: 452, Duden Grammatik 2009: 161) if expressed as the constraint $NON\text{-}GER. RIVERS\ IN\text{-}E \rightarrow *M, *N$, would be concealing the ranking $-E \rightarrow *M, *N \gg NON\text{-}GER. RIVERS \rightarrow *F, *N$. Therefore, in such cases, rules have been separated into two constraints, in this case: $NON\text{-}GER. RIVERS \rightarrow *F, *N$ and $-E \rightarrow *M, *N$.³

It should also be noted that there are a number of gender assignment principles in the literature which contradict each other, e.g. “musical instruments are neuter” (e.g. Steinmetz 1986: 197, Nelson 1998: 219) vs. “musical instruments are feminine” (e.g. Mills 1986: 27), and “generic terms for people are masculine” (e.g. Köpcke 1982: 72, Mills 1986: 27, Hoberg 2004: 103, Chan 2005: 91ff., Onysko 2007: 159) vs. “generic terms for people are neuter” (e.g. Flämig 1991: 456). In such cases, all versions of the constraint will be tested (e.g. both $GEN. PERSON \rightarrow *F, *N$ and $GEN. PERSON \rightarrow *M, *F$), and it will be seen which is most plausible according to the criteria (see 2.9).

In total, 378 constraints were found in the literature – 159 based on semantic features and 219 based on the formal properties of a noun. However, for the purposes of this study, it is not necessary to assess each of the 378 constraints using Enger’s (2009) six criteria, since not all of these constraints

³ Other cases include e.g. “nouns denoting groups in *-a*” (Nelson 1998: 214) and “nouns in *-ent* denoting people” (Durrell 2011: 10). N.B. Constraints such as $-LT\ MONO. \rightarrow *M, *F$ “monosyllables in [It] cannot be masculine or feminine” are not an example of this, since they could, for instance, be alternatively phrased as $(C)(C)(C)(V)(V)LT \rightarrow *F, *N$.

are applicable to the sample of 592 nouns being tested and therefore will not be relevant to the analysis in Chapter 3.

Determining which of the constraints are applicable to the noun sample is largely straightforward. For most constraints, it is a simple case of finding at least one noun in the sample with a particular formal feature or a particular meaning. For example, it is clear that the constraint $-UNG \rightarrow *M, *N$ is relevant to the sample due to the presence of the nouns *Bloßstellung* and *Verzögerung*, yet that the constraint $MOTORBIKES \rightarrow *M, *N$ is not, as there are simply no nouns denoting motorbikes in the sample. However, there are a small number of semantic constraints whose vagueness means that it is unclear whether they are relevant to the sample or not. For instance, might $FUNCTIONAL\ HOLLOWES \rightarrow *M, *F$ be applicable to *Aufzugsschacht* ‘lift shaft’ or *Orgelpfeife* ‘organ pipe’? Could it be argued that $SEMANTIC\ DIMINUTIVES \rightarrow *M, *N$ is applicable to *Rinne* ‘groove’ in light of nouns such as *Schlucht* ‘ravine’? In such cases, where there is any possibility at all of a constraint applying to one or more nouns in the sample, the constraint will be assessed by the criteria. Only in cases where it is clear that they have no possible application to the sample have constraints been categorised as irrelevant.

Consequently, 237 out of the total 378 constraints will be tested in terms of their validity using Enger’s (2009) six criteria.

	All proposed	Relevant/potentially relevant to sample	Not relevant to sample
Meaning-based	159	94	65
Form-based	219	143	76
Total	378	237	141

See Appendix B for the full list of constraints found in the literature, along with examples of relevant or potentially relevant nouns from the sample where applicable.

It is important to highlight that evaluating just 237 out of the total 378 proposed constraints will not affect the outcome of the analysis conducted in Chapter 3, since the 141 unevaluated constraints bear no relevance to the nouns in the sample to be tested. Technically, of course, all valid constraints – even those which are not applicable to the noun being tested – are present in an OGAT tableau, i.e. a native speaker’s gender assignment system, at all times. For practical reasons alone, however, only the constraints which are relevant to the nouns being tested in Chapter 3 will be assessed for their plausibility.

Below, each of the 237 constraints will be evaluated according to Enger's (2009) six criteria. The discussion will address exactly how the fulfilment of each criterion is measured, the methods involved in assessing the constraints and a final judgement with regard to each constraint's plausibility.

2.3 Criterion 1: Coverage of candidate nouns

'A plausible rule should cover a reasonable share of the possible candidate nouns' (Enger 2009: 1290)

The first of Enger's (2009) criteria stipulates that a gender assignment principle must be true for a reasonable share of the nouns to which it applies in order to be plausible. It holds that if a gender assignment principle covers all possible candidate nouns, i.e. is exceptionless, then native speakers have a higher chance of making the association between the relevant feature and gender, and that the principle would be of practical use for the speaker and therefore more likely to be valid. If a principle has a high proportion of exceptions, then not only would it be more difficult to pick up on the association, but the principle would be less useful, as a speaker would have to learn both the principle and a long list of exceptions.

Moreover, the coverage of a high proportion of candidate nouns is the justification given for most of the gender assignment principles proposed in the literature (such as those listed in 1.3). Most of the rules outlined in e.g. Zubin & Köpcke (1981, 1984, 1986), Köpcke (1982), Mills (1986), Wegener (1995), Nelson (1998) etc. are justified by their lack of exceptions. Additionally, a number of linguists (e.g. Mills 1986: 114, Heringer 1995: 212, Wegener 2000: 514) explicitly state that the "validity" and "clarity" of gender assignment rules are contingent upon the number of exceptions they allow.

The first criterion therefore seems key in determining the plausibility of a gender assignment principle. However, Enger (2009) does not provide a method by which to assess principles according to this criterion, specifically leaving open the question of what might constitute a 'reasonable share' of candidate nouns. He does, however, state that principles which cover 'almost all' candidate nouns with just one or two exceptions would satisfy this criterion, but those with as many exceptions as examples or more exceptions than examples would not (Enger 2009: 1290f.).

It seems reasonable to suggest that a constraint is useful if it is more effective in predicting a noun's gender than if it were left down to chance. Since there are three genders in German, it could be argued that a "reasonable share" is anything greater than 33.3%. Alternatively, the distribution of nouns across the three genders could be taken into account. However, this would not be straightforward,

since estimates of the distribution of nouns across genders vary considerably (see Figure 3 in 1.4.5.3). Regardless of whether noun distribution is taken into account or not, a key problem with this approach is that the “reasonable share” threshold would be low. It would allow constraints which have more exceptions than examples (i.e. those which cover less than 50% of candidate nouns) to satisfy the criterion, which directly goes against Enger’s (2009: 1290f.) statement that constraints with as many exceptions as examples or more exceptions than examples do not meet the criterion.

An alternative approach to determining what constitutes a “reasonable share” would be to make use of Yang’s (2005) Tolerance Principle. Yang (2005) proposes that it is possible to calculate the number of exceptions a productive⁴ rule can tolerate before it becomes less computationally efficient to generalise a rule than to individually store each lexical item. Yang (2005) bases his model on psycholinguistic evidence relating to processing times. He reasons that when a speaker is confronted with a possible rule, for instance ‘nouns ending in schwa are feminine’, and encounters both nouns which adhere to this rule and nouns which do not, the speaker has two options: either they may store each lexical item individually in a list (ordered by frequency) and search the list each time they wish to produce a noun ending in a schwa in order to find its gender, or the speaker generalises a productive rule from the data and stores only the exceptions to the rule in a list. Each time the speaker wishes to produce a noun ending in a schwa, they would then first search the list of exceptions for the target noun, and if it is not there, apply the rule.

Yang’s (2005) Tolerance Principle works on the assumption that a speaker will always employ the faster method. That is to say, a productive rule will only be created if it is more computationally efficient to do so. Therefore, given that the more exceptions a rule has, the longer it takes to process the rule, Yang (2005) argues that once the number of exceptions reaches a certain threshold, it then becomes more efficient to abandon the rule and store all items individually instead. Formally, he states that: if there is a linguistic rule or process R which in theory could apply to a set of words N , and of these there are M exceptions which do not follow R , then:

$$(1) \text{ Tolerance Principle: } R \text{ can only be productive iff } M < \frac{N}{\ln N}$$

That is to say, a linguistic process is only productive if and only if the number of exceptions (M) is less than the total number of words the process applies to (N) divided by the natural logarithm (\ln) of N . Accordingly, the threshold at which a process or rule ceases to be efficient (M_c) can be expressed in

⁴ Yang (2005: 272) defines “productive” as ‘predictable’ and ‘generalizable’. He states that ‘a rule is productive if it automatically applies to a set of lexical items characterized by certain properties, producing predictable derived forms, and can extend to others, including novel items that have the same properties’.

terms of N : $M_c \approx N/\ln N$. This would mean that, for example, if N were equal to 100, then the maximum number of exceptions R could tolerate (M_c) would be 22. In terms of criterion 1, then, the Tolerance Principle would dictate that, in this case, a “reasonable share” would be approximately equal to 78%. For a full account of the derivation of the formula see Yang (2005).

The Tolerance Principle has support from language acquisition data, e.g. Schuler et al. (2016), but it has also not gone uncriticised. For instance, O’Donnell (2015: 228ff.) lists three potential problems with Yang’s (2005) model, one being, for example, that it cannot account for empirical evidence which suggests that certain highly-frequent word forms containing regular and productive suffixes are stored. However, O’Donnell (2015: 228ff.) concedes that Yang’s (2005) model could be extended to account for some of the issues that he raises.

In any case, the Tolerance Principle will prove useful for helping to quantify a “reasonable share”, given that calculating a threshold using a model based on psycholinguistic evidence which also has support from language acquisition data is infinitely better than deciding on a threshold at random, even if the model might have imperfections.

Method

In order to use the Tolerance Principle to determine what a “reasonable share” might be, then the threshold of exceptions (M_c) and the number of exceptions (M) must be calculated for each constraint. For this to be done, samples of possible candidate nouns need to be collected for each constraint, so that it can be seen whether the actual number of exceptions found in the samples is above or below the threshold, which will be calculated according to the size of the sample (N).

For constraints relating to the form of a noun, the majority of the samples have been collected using the Duden-Bibliothek Express software (an easily searchable version of the Duden), which allows all nouns in the Duden with a particular formal feature to be found. For some of the constraints involving monosyllabic nouns, Köpcke’s (1982) sample of 1466 monosyllables (also originally from the Duden) has been used.

For constraints relating to semantic properties of nouns, samples have been collected from various sources such as: the Oxford-Duden pictorial dictionary (1994), the Duden Online dictionary (2018), the Duden thesaurus (2010) and a variety of websites (see Appendix C for the full list of sources). For some of the more unclear constraints (i.e. those which fare badly by the fourth criterion – see 2.6), it was naturally more difficult to collect the noun samples, since personal judgement had to be used to decide which nouns belonged to the category and which did not. This was the case with constraints

such as F.HOLLOWS→*M,*F ‘functional hollows cannot be masculine or feminine’ and SHEET/STRIP/SLAB→*M,*F ‘nouns denoting sheets, strips or slabs cannot be masculine or feminine’. Five of the constraints invoke semantic categories which are so vague that it was simply not possible to gather a sample of possible candidate nouns, including IMPROPER SUPERORDINATES→*M,*F; SEMANTIC DIMINUTIVES→*M,*N; and SUPERORDINATES OF INDIRECT REFERENCE→*M,*F.

Results

Table 1 shows all of the constraints tested along with values for the total number of candidate nouns in the sample tested (N), the number of exceptions found in the sample (M), the exceptions threshold as predicted by the Tolerance Principle (M_c) and an indication of whether the constraint is valid based on these figures (i.e. whether $M < M_c$ or not). The table is ordered according to the percentage coverage of the constraints for perspective.

Constraint (R)	Total candidate nouns in sample (N)	Exceptions (M)	Exceptions tolerated (M_c)	Valid?	% Coverage ↓
MUS. NOTES→*M,*F	21	0	6.90	✓	100.00
CHEESES→*F,*N	34	0	9.64	✓	100.00
LETTERS→*M,*F	56	0	13.91	✓	100.00
CITIES/TOWNS→*M,*F	60	0	14.65	✓	100.00
METALS→*M,*F	88	0	19.65	✓	100.00
LANGUAGES→*M,*F	107	0	22.90	✓	100.00
NAME: MAN→*F,*N	555	0	87.83	✓	100.00
NAME: WOMAN→*M,*N	514	0	82.34	✓	100.00
-ISMUS→*F,*N	483	0	78.16	✓	100.00
-OR→*F,*N	245	0	44.54	✓	100.00
-(O)LOGE→*F,*N	94	0	20.69	✓	100.00
-ITIS→*M,*N	73	0	17.01	✓	100.00
-UR SUFF.→*M,*N	66	0	15.75	✓	100.00
-INER→*F,*N	20	0	6.68	✓	100.00
-THEK→*M,*N	17	0	6.00	✓	100.00
-SAL→*M	15	0	5.54	✓	100.00
VCCC MONO→*M	4	0	2.89	✓	100.00
-ANER→*F,*N	116	0	24.40	✓	100.00
-EDER→*M,*F	12	0	4.83	✓	100.00
-EKT→*F	22	0	7.12	✓	100.00
-SKOP→*M,*F	32	0	9.23	✓	100.00
-NIS→*M	40	0	10.84	✓	100.00
-ON.UNSTRESS.→*M,*F	54	0	13.54	✓	100.00
-ANZ→*M,*N	57	0	14.10	✓	100.00
VER- NO SUFF.→*F	59	0	14.47	✓	100.00
-MA→*M,*F	60	0	14.64	✓	100.00
-AGE→*M,*N	69	0	16.30	✓	100.00
-IV→*F	69	0	16.30	✓	100.00
-ADE→*M,*N	72	0	16.84	✓	100.00
-IE [iə]→*M,*N	83	0	18.78	✓	100.00
-MENT→*M,*F	98	0	21.37	✓	100.00
-LING→*F,*N	105	0	22.56	✓	100.00
-TUM→*F	118	0	24.73	✓	100.00
-CHEN→*M,*F	126	0	26.05	✓	100.00
-ENZ→*M,*N	130	0	26.71	✓	100.00

-IT→*F	165	0	32.32	✓	100.00
CONVERSION→*M,*F	182	0	34.97	✓	100.00
-UM→*M,*F	235	0	43.04	✓	100.00
-IUM→*M,*F	260	0	46.76	✓	100.00
-IST→*F,*N	360	0	61.16	✓	100.00
-SCHAFT→*M,*N	380	0	63.97	✓	100.00
-EI→*M,*N	384	0	64.53	✓	100.00
-TÄT→*M,*N	439	0	72.15	✓	100.00
-IE ['i:]→*M,*N	788	0	118.15	✓	100.00
-ION→*M,*N	913	0	133.94	✓	100.00
-UNG→*M,*N	1074	0	153.89	✓	100.00
-HEIT/-(IG)KEIT→*M,*N	1579	0	214.41	✓	100.00
-ER SUFF.→*F,*N	1796	0	239.68	✓	100.00
-IN SUFF.→*M,*N	3155	0	391.60	✓	100.00
-ING→*M,*F	165	1	32.32	✓	99.39
LEATHER TYPES→*M,*F	146	2	29.30	✓	98.63
-EN→*F	435	6	71.60	✓	98.62
-ANT→*F,*N	137	2	27.85	✓	98.54
-TUM→*M,*F	118	3	24.73	✓	97.46
-IK→*M,*N	319	9	55.33	✓	97.18
D- MONO→*F	35	1	9.84	✓	97.14
-ANG→*F,*N	59	2	14.84	✓	96.72
GEN. PERSON→*F,*N	544	18	86.36	✓	96.69
CCVC MONO→*F,*N	29	1	8.61	✓	96.55
-US→*F,*N	643	24	99.44	✓	96.27
MALE ANIMALS→*F,*N	26	1	7.98	✓	96.15
-EUR/ÖR→*N	117	5	24.57	✓	95.73
-ETTE→*M,*N	90	4	20.00	✓	95.56
PRIMATES→*F,*N	20	1	6.68	✓	95.00
V _{STRESSED} [m]→*M,*F	88	4	19.65	✓	95.45
ELEMENTS→*M,*F	118	6	24.73	✓	94.92
-ETT→*M,*F	74	4	17.19	✓	94.59
-INE→*M,*N	88	5	19.65	✓	94.32
-B→*F	122	7	25.40	✓	94.26
R- MONO→*F	64	4	15.39	✓	93.75
ROCKS/MINERALS→*F,*N	144	9	28.97	✓	93.75
-L MONO→*F	79	5	18.08	✓	93.67
-ILLE→*M,*N	47	3	12.21	✓	93.62
-EUR/ÖR→*F,*N	117	8	24.57	✓	93.32
∅V _{LONG} MONO→*F	13	1	5.07	✓	92.31
YOUNG ANIMALS→*M,*F	26	2	7.98	✓	92.31
-ER→*F,*N	2245	173	290.94	✓	92.29
FRUITS→*M,*N	62	5	15.02	✓	91.94
-C S MONO→*F,*N	160	13	31.53	✓	91.88
T- MONO→*F	49	4	12.59	✓	91.84
VV _{diphth} C MONO→*F	134	11	27.36	✓	91.79
-IT→*F,*N	165	14	32.32	✓	91.52
COUNTRIES→*M,*F	182	16	34.97	✓	91.21
CCVV _{DIPHTHONG} - MONO→*F	67	6	14.84	✓	91.04
BODIES OF WATER→*N	32	3	9.23	✓	90.63
ALC. DRINKS→*F,*N	62	6	15.02	✓	90.32
-UR STRESSED→*M,*N	129	13	26.54	✓	89.92
-ST/-SCHT/-ZT→*N	89	9	19.83	✓	89.89
MUS. INSTR.→*M	91	10	20.17	✓	89.01
CVC MONO→*F	403	47	67.18	✓	88.34
TREES→*M,*N	25	3	7.77	✓	88.00
-ENT→*F,*N	91	11	20.17	✓	87.91
RIVERS NON-EUROPE→*F,*N	221	27	40.94	✓	87.33

[j]C MONO→*F,*N	169	22	32.94	✓	86.98
DEVERBAL -T→*M,*N	22	3	7.12	✓	86.36
V _{LONG} C MONO→*F	307	43	53.61	✓	85.99
ACAD. DISCIPLINES→*M,*N	106	15	22.73	✓	85.85
DEVERBAL STEM→*F,*N	400	57	66.76	✓	85.75
MUSIC TYPES→*F,*N	56	8	13.91	✓	85.71
V _{SHORTSTOP} C→*F,*N	124	18	25.72	✓	85.48
V _{LONG} MONO→*F	369	54	62.43	✓	85.37
INTROVERSION→*M,*N	34	5	9.64	✓	85.29
MON. UNITS→*F,*N	75	12	17.37	✓	84.00
-IVE→*M,*N	25	4	7.77	✓	84.00
SUP.→*M,*F	56	9	13.91	✓	83.93
MEN→*F,*N	73	12	17.01	✓	83.56
ANG. SHAPES→*M,*F	36	6	10.05	✓	83.33
DR→*F,*N	24	4	7.55	✓	83.33
SEASONS→*F,*N	6	1	3.35	✓	83.33
TR→*F,*N	47	8	12.21	✓	82.98
V _{STRESSED} [l]→*M,*F	70	12	16.48	✓	82.86
SCI. UNITS→*M,*F	29	5	8.61	✓	82.76
CVCC MONO→*N	354	60	60.31	✓	82.20
CVCC MONO→*F	354	60	60.31	✓	82.20
RIVERS NON-GER.→*F,*N	84	15	18.96	✓	82.14
WOMEN→*M,*N	138	25	28.01	✓	81.88
-/tj/;/Ntj/;/Nj/ MONO→*F,*N	32	6	9.23	✓	81.25
GAMES→*M,*F	79	15	18.08	✓	81.01
-(C)(C)// MONO→*F,*N	84	16	18.96	✓	80.95
-IG/-ICH→*F,*N	42	8	11.24	✓	80.95
MOUNTAINS→*F,*N	81	16	18.43	✓	80.25
-AT→*M,*F	186	38	35.59	x	79.57
-IN [i:n]→*M,*F	181	37	34.82	x	79.56
-E→*M,*N	3407	735	418.88	x	78.43
DAYS→*F,*N	36	8	10.05	✓	77.78
FABRIC→*F,*N	45	10	11.82	✓	77.78
SUFFIX.-E→*M,*N	1943	432	256.50	x	77.76
FISH→*F,*N	51	12	12.97	✓	76.47
FLOWERS→*M,*N	38	9	10.45	✓	76.32
-IS→*M,*N	198	49	37.44	x	75.25
-A→*M,*N	771	191	115.98	x	75.22
CCVC MONO→*F,*N	298	75	52.31	x	74.83
GE→*M,*F	607	155	94.72	x	74.46
CCVCC→*F,*N	152	39	30.26	x	74.34
REGIONS→*M,*F	210	54	39.27	x	74.29
CHEMICAL COMP./SUBST.→*M,*F	113	30	23.90	x	73.45
-N(C)(C) MONO→*F,*N	290	78	51.15	x	73.10
ISLANDS→*M,*F	120	33	25.07	x	72.50
-/r/ STOP (C) MONO→*F,*N	65	18	15.57	x	72.31
SPEECH→*M,*N	43	12	11.43	x	72.09
FEAR→*M,*N	32	9	9.23	✓	71.88
-IV STRESSED→*M,*F	49	14	12.59	x	71.43
-AN→*M,*F	135	39	27.52	x	71.11
-AR→*M,*F	54	16	13.54	x	70.37
GROUPS→*M,*N	74	22	17.19	x	70.27
-RIS→*M,*N	10	3	4.34	✓	70.00
-IV→*M,*F	69	21	16.30	x	69.57
VEG. ROOT→*M,*N	16	5	5.77	✓	68.75
NO INFO.→*M,*F	32	10	9.23	x	68.75
CONDUCT→*M,*N	54	17	13.54	x	68.52
-[l]C MONO→*F,*N	89	29	19.83	x	67.42

CVCCC MONO→*F,*N	55	18	13.72	x	67.27
-ON→*M,*F	243	80	44.24	x	67.08
TIMES OF DAY→*F,*N	6	2	3.35	✓	66.67
-ÜR, -ÜHR /y:r/→*M,*N	6	2	3.35	✓	66.67
TEMPERATURES→*M,*N	9	3	4.10	✓	66.67
-FT/-CHT→*M,*N	73	25	17.01	x	65.75
INSECTS→*M,*N	32	11	9.23	x	65.63
DISEASES→*M,*N	95	33	20.86	x	65.26
-ST/-SCHT/-ZT→*F,*N	89	31	19.83	x	65.17
POWER/STRENGTH→*M,*N	34	12	9.64	x	64.71
MONOSYLL.→*F,*N	1466	528	201.09	x	63.98
BIRDS→*F,*N	69	25	16.30	x	63.77
-EKT→*F,*N	22	8	7.12	x	63.64
KR- MONO→*F,*N	33	12	9.44	x	63.64
CV MONO→*N	39	11	10.65	x	63.64
VC MONO→*M,*F	22	8	7.12	x	63.64
WASTE→*F,*N	30	11	8.82	x	63.33
VCC MONO→*F,*N	27	10	8.19	x	62.96
MESH→*M,*F	29	11	8.61	x	62.07
MAMMALS→*F,*N	68	27	16.12	x	60.29
-SAL→*M,*F	15	6	5.54	x	60.00
-AR SUFF.→*F,*N	47	19	12.21	x	59.57
STICK→*F,*N	56	23	13.91	x	58.93
PICTURES→*M,*F	33	14	9.44	x	57.58
-IV UNSTRESSED→*F,*N	28	12	8.40	x	57.14
RIVERS EUROPE→*M,*N	105	46	22.56	x	56.19
-NIS→*M,*F	40	18	10.84	x	55.00
MUS. INSTR.→*M,*N	91	41	20.17	x	54.95
-AL→*M,*F	148	69	29.62	x	53.38
-IG→*F,*N	15	7	5.54	x	53.33
-EHR/-ÄHR→*M,*N	17	8	6.00	x	52.94
BOATS→*F,*N	102	48	22.05	x	52.94
DOM. ANIMALS→*M,*F	25	12	7.77	x	52.00
-EN→*M,*F	435	209	71.60	x	51.95
ELAB. INSTR.→*M,*N	48	24	12.40	x	50.00
WORD CLASSES→*M,*F	12	6	4.83	x	50.00
UNITS OF TIME→*F,*N	45	23	11.82	x	48.89
-O→*M,*F	386	198	64.81	x	48.70
NAUTICAL→*M,*N	126	65	26.05	x	48.41
MUS. TERMS→*M,*F	151	78	30.10	x	48.34
DWELL→*M,*F	50	26	12.78	x	48.00
SHEET/STRIP/SLAB→*M,*F	42	22	11.24	x	47.62
-EN→*F,*N	435	232	71.60	x	46.67
-EL→*F,*N	596	324	93.27	x	45.64
TIME: SHORT→*F,*N	11	6	4.59	x	45.45
-NIS→*M,*N	40	22	10.84	x	45.00
MUS. COMP.→*M,*F	126	70	26.05	x	44.44
PREP. DISHES→*M,*F	94	53	20.69	x	43.62
EXTROVERSION→*F,*N	35	20	9.84	x	42.86
-EE/-ÉE/-É→*M,*F	132	76	27.03	x	42.42
LIVING/WORKING→*M,*F	92	54	20.35	x	41.30
-IER /i:r/ POLY.→*M,*F	39	23	10.65	x	41.03
GESTURES→*M,*N	27	16	8.19	x	40.74
-AR SUFF.→*M,*F	47	28	12.21	x	40.43
-SAL→*M,*N	15	9	5.54	x	40.00
PLANTS→*F,*N	99	60	21.54	x	39.39
SHARP→*M,*N	39	24	10.65	x	38.46
PCC GARMENT→*M,*F	42	26	11.24	x	38.10

INSTITUTIONS→*M,*F	45	28	11.82	x	37.78
-EKT→*M,*F	22	14	7.12	x	36.36
MUS. INSTR.→*M,*F	91	60	20.17	x	34.07
-EE/-ÉE/-É→*F,*N	132	89	27.03	x	32.58
ANNOUNCE→*M,*F	43	29	11.43	x	32.56
HUNTING→*M,*N	100	69	21.71	x	31.00
OPEN WATER→*M,*N	23	16	7.34	x	30.44
-IV→*F,*N	69	48	16.30	x	30.43
ORGANS→*M,*F	17	12	6.00	x	29.41
-LT MONO→*M,*F	32	23	9.23	x	28.13
-EL→*M,*N	596	433	93.27	x	27.35
-EL→*M,*F	596	435	93.27	x	27.01
F. HOLLOWES→*M,*F	56	42	13.91	x	25.00
-EE/-ÉE/-É→*M,*N	132	99	27.03	x	25.00
OPENINGS→*M,*F	65	49	15.57	x	24.62
VER→*F	709	561	108.02	x	20.87
-AT→*F,*N	186	148	35.59	x	20.43
GROUPS→*M,*F	74	60	17.19	x	18.92
-ENT→*M,*F	91	81	20.17	x	10.99
-IT→*M,*F	165	151	32.32	x	8.48
-ER→*M,*F	2245	2138	290.94	x	4.77
-ER→*M,*N	2245	2179	290.94	x	2.94
-TUM→*F,*N	118	115	24.73	x	2.54
GEN. PERSON→*M,*F	544	532	86.36	x	2.21
IMPROPER SUPERORDINATES→*M,*F	–	–	–	–	–
INTEGRATED PARTS→*F,*N	–	–	–	–	–
SEMANTIC DIMINUTIVES→*M,*N	–	–	–	–	–
SEPARABLE PARTS→*M,*F	–	–	–	–	–
SUPERORDINATES OF INDIRECT REFERENCE→*M,*F	–	–	–	–	–

Table 1: Assessment of constraints using Yang's (2005) Tolerance Principle

Table 1 shows that all tested constraints with a coverage of 80.25% or higher are deemed valid according to Yang's (2005) model, and all tested constraints with a coverage of 65.75% or lower are deemed not to cover a "reasonable share". The middle section containing constraints with a coverage of between 66.67% and 79.57%, however, merits further consideration, because there are some constraints in this section whose M value is only marginally higher than their Mc value. Since many of these constraints also have a relatively high percentage coverage, it is worth looking at these again.

It is important to recall that the values for N and M have been calculated on the basis of the samples of nouns collected. Rather than N being the actual number of all nouns to which R could potentially apply, it is instead the number of candidate nouns found in the sample. We should therefore allow for some margin of error, given the imperfect nature of sampling. Furthermore, Yang (2005) himself states that the tolerance threshold for exceptions (Mc) is only *roughly equal* to $N/\ln N$. Some flexibility in the calculations is therefore necessary.

If we were to allow for a margin of error for M_c of, for instance, 10% of the total candidate nouns in the sample (N), this would enable M_c to serve as a rough guide rather than a strict limit. Table 2 shows the results produced for the middle section of Table 1 when using a 10% margin.

Constraint (R)	Total candidate nouns in sample (N)	Exceptions (M)	Exceptions tolerated (M_c)	$M_c + 10\%$ of N	Valid?	% Coverage ↓
-AT→*M,*F	186	38	35.59	54.19	✓	79.57
-IN [i:n]→*M,*F	181	37	34.82	52.92	✓	79.56
-E→*M,*N	3407	735	418.88	759.58	✓	78.43
DAYS→*F,*N	36	8	10.05	13.65	✓	77.78
FABRIC→*F,*N	45	10	11.82	16.32	✓	77.78
SUFFIX.-E→*M,*N	1943	432	256.50	450.90	✓	77.76
FISH→*F,*N	51	12	12.97	18.07	✓	76.47
FLOWERS→*M,*N	38	9	10.45	14.25	✓	76.32
-IS→*M,*N	198	49	37.44	57.24	✓	75.25
-A→*M,*N	771	191	115.98	193.08	✓	75.22
CCVC MONO→*F,*N	298	75	52.31	82.11	✓	74.83
GE→*M,*F	607	155	94.72	155.42	✓	74.46
CCVCC→*F,*N	152	39	30.26	45.36	✓	74.34
REGIONS→*M,*F	210	54	39.27	60.27	✓	74.29
CHEM. COMP./SUBST.→*M,*F	113	30	23.90	35.20	✓	73.45
-N(C)(C) MONO→*F,*N	290	78	51.15	80.15	✓	73.10
ISLANDS→*M,*F	120	33	25.07	37.07	✓	72.50
-/r/ STOP (C) MONO→*F,*N	65	18	15.57	22.07	✓	72.31
SPEECH→*M,*N	43	12	11.43	15.73	✓	72.09
FEAR→*M,*N	32	9	9.23	12.43	✓	71.88
-IV STRESSED→*M,*F	49	14	12.59	17.49	✓	71.43
-AN→*M,*F	135	39	27.52	41.02	✓	71.11
-AR→*M,*F	54	16	13.54	18.94	✓	70.37
GROUPS→*M,*N	74	22	17.19	24.59	✓	70.27
-RIS→*M,*N	10	3	4.34	5.34	✓	70.00
-IV→*M,*F	69	21	16.30	23.20	✓	69.57
VEG. ROOT→*M,*N	16	5	5.77	7.37	✓	68.75
NO INFO.→*M,*F	32	10	9.23	12.43	✓	68.75
CONDUCT→*M,*N	54	17	13.54	18.94	✓	68.52
-[l]C MONO→*F,*N	89	29	19.83	28.73	x	67.42
CVCCC MONO→*F,*N	55	18	13.72	19.22	✓	67.27
-ON→*M,*F	243	80	44.24	68.54	x	67.08
TIMES OF DAY→*F,*N	6	2	3.35	3.95	✓	66.67
-ÜR,-ÜHR /y:r/→*M,*N	6	2	3.35	3.95	✓	66.67
TEMPERATURES→*M,*N	9	3	4.10	5	✓	66.67
-FT/-CHT→*M,*N	73	25	17.01	24.31	x	65.75
INSECTS→*M,*N	32	11	9.23	12.43	✓	65.63
DISEASES→*M,*N	95	33	20.86	30.36	x	65.26
-ST/-SCHT/-ZT→*F,*N	89	31	19.83	28.73	x	65.17
POWER/STRENGTH→*M,*N	34	12	9.64	13.04	✓	64.71
MONOSYLL.→*F,*N	1466	528	201.09	347.60	x	63.98
BIRDS→*F,*N	69	25	16.30	23.20	x	63.77
-EKT→*F,*N	22	8	7.12	9.32	✓	63.64
KR- MONO→*F,*N	33	12	9.44	12.74	✓	63.64
CV MONO→*N	39	11	10.65	14.55	✓	63.64
VC MONO→*M,*F	22	8	7.12	9.32	✓	63.64
WASTE→*F,*N	30	11	8.82	11.82	✓	63.33
VCC MONO→*F,*N	27	10	8.19	10.89	✓	62.96
MESH→*M,*F	29	11	8.61	11.51	✓	62.07

-SAL→*M,*F	15	6	5.54	7.04	✓	60.00
-AR SUFF.→*F,*N	47	19	12.21	16.91	x	59.57
STICK→*F,*N	56	23	13.91	19.51	x	58.93
PICTURES→*M,*F	33	14	9.44	12.74	x	57.58
					x	
-IG→*F,*N	15	7	5.54	7.04	✓	53.33
WORD CLASSES→*M,*F	12	6	4.83	6.03	✓	50.00
UNITS OF TIME→*F,*N	45	23	11.82	16.32	x	48.89
-O→*M,*F	386	198	64.81	103.41	x	48.70
					x	

Table 2: Assessment of constraints where $M_c = M_c + 10\%N$

With a 10% margin, all constraints with a coverage of 68.52% or above are considered to cover a “reasonable share”, and almost all constraints with a coverage below 60.00% are not. This is in line with Enger’s (2009: 1290f.) requirement for a “reasonable share” to cover more than half of the candidate nouns.

There are, however, a small number of outliers in Table 2. The first set are the seven constraints which have relatively high percentage coverages of between 67.42% and 63.77%, yet are considered to have too many exceptions according to Yang’s (2005) model. The other outliers are -IG→*F,*N and WORD CLASSES→*M,*F, which both have relatively low percentage coverages, yet are deemed valid. The reason for these outliers is related to the sizes of the samples (N) used to evaluate the constraints. M_c is calculated as being roughly equal to $N/\ln N$, where \ln is a function used to determine the amount of time needed for a certain level of exponential growth to be reached. This means that the effects of the Tolerance Principle are greater the larger that N becomes. Specifically, when N is small, M_c will be large relative to N , and when N is large, M_c will be small relative to N . For instance, if N is 10, then M_c would be 4.34, which is almost half of N . However, if N is 1000, M_c would be 144.76, which is only around one seventh of N . Presumably the reasoning behind this is that it would take considerably longer to process half of N when N is 1000 than to process half of N if N is only 10. For this reason, M_c is calculated in terms of N rather than as a blanket percentage threshold.

This explains why the set of seven constraints are deemed invalid despite having high percentage coverages, since the sample sizes are large relative to that of other constraints (69 – 1466) and therefore M_c is small in relation to these figures. By the same token, -IG→*F,*N and WORD CLASSES→*M,*F are considered valid in spite of their low percentage coverages because the sample sizes are comparatively small (15 and 12) and thus M_c is large relative to these figures.

In sum, then, Enger’s (2009) first criterion is key in determining the plausibility of a gender assignment principle, since, if a principle is unable to account for a “reasonable share” of candidate nouns, it is of little use to native speakers. A promising solution to the challenge of quantifying a “reasonable share”

comes in the form of Yang's (2005) Tolerance Principle, which provides a simple yet effective method for determining what might constitute a "reasonable share" for an OGAT constraint. Used with a margin of error due to the approximate nature of the Tolerance Principle and also in order to compensate for the imperfect nature of sampling, the Tolerance Principle produces plausible results and is able to serve as a guide for determining which constraints are valid on the basis of the proportion of candidate nouns they account for.

2.4 Criterion 2: Typological parallels

'If a rule has parallels in other languages, that is an advantage' (Enger 2009: 1291)

Enger's second criterion seeks typological justification for constraints, working on the basis that a claim about the relevance of a particular feature to gender assignment in one language is made more plausible by the relevance of the same feature to gender assignment in other languages. In standard OT, typological support for constraints is highly valued, given that constraints are argued to be universal (e.g. Prince & Smolensky 1993/2004, Kager 1999: 4, McCarthy 2002: 108, 2008: 235).⁵

Indeed, Hayes (2004: 291) identifies typological evidence as the most popular way to justify a constraint in standard OT, stating that a valid constraint is one that works in many languages. Additionally, McCarthy (2008: 212) considers typological justification to be the most compelling kind of justification for an OT constraint, and Kager (1999: 1) even remarks that 'we should be very careful about positing any constraint lacking [...] typological motivation'.

Typological support for constraints is therefore extremely valuable. The main challenge with this criterion, however, as with all of Enger's (2009) criteria, is how exactly fulfilment of the criterion can be determined.

Enger (2009: 1291) gives three examples of gender assignment principles which he considers to have 'clear typological parallel[s]' and thus satisfy the criterion. The first is the Norwegian principle "trees are feminine" (Trosterud 2001: 41), which has parallels in Italian, in which trees are typically masculine (Thornton 2009) and Latin, in which trees are typically feminine (Enger 2009: 1291). There is also a parallel to this principle in German, in which trees are typically feminine (e.g. Flämig 1991: 452, Hickey 2000: 629, Menzel 2004: 63, Thomoglou 2004: 34, Chan 2005: 95, Weinrich 2007: 328, Engel 2009: 272). The second principle is "cars are feminine", proposed for Italian by Thornton (2009: 25), which

⁵ Note that for OGAT, only the presence of GENDER FEATURES and a default hierarchy is bound by this; individual GENDER FEATURES constraints are language specific (Rice 2006: 1410) and are therefore not required by OT to have typological justification. However, if they do have typological parallels, then Enger (2009: 1291) argues that this is compelling evidence for their existence.

has parallels in Norwegian (Enger 2009: 1291) and in German, in which cars are typically masculine (e.g. Flämig 1991: 452, Hickey 2000: 630, Menzel 2004: 63, Hoberg 2004: 108, Köpcke & Zubin 2005, Weinrich 2007: 328, Duden Grammatik 2009: 162, Engel 2009: 272). The third principle considered to meet this criterion is the German principle “alcoholic drinks are masculine” (e.g. Zubin & Köpcke 1981: 444, Köpcke 1982: 72, Zubin & Köpcke 1984: 44, Köpcke & Zubin 1996: 479, Hickey 2000: 630, Menzel 2004: 63, Thomoglou 2004: 34, Chan 2005: 94, Engel 2009: 272, Duden Grammatik 2009: 159, Eisenberg 2013: 138), which has a parallel in Norwegian (Enger 2009: 1291).

Two key facts can be deduced from these claims made by Enger (2009: 1291), namely that, firstly, typological evidence from just one other language is deemed to be sufficient for a constraint to fulfil this criterion, and secondly, that the feature need not be associated with the same gender in other languages.

Enger (2009: 1291) also states that ‘we should not rely on an over-literal reading of the criterion’, suggesting that the typological parallels need not always be entirely direct. For instance, Enger (2009: 1291) observes that, for obvious pragmatic reasons, the principle “names of towns on the left bank of the Volga are feminine and names of towns on the right bank are masculine in Russian before 1917” (Superanskaja 1965: 58, cited in Corbett & Fraser 2000: 321) is unlikely to have direct parallels in many languages other than Russian (and Ukrainian). However, he states that the German principle of “non-German rivers are feminine” (e.g. Flämig 1991: 451, Duden Grammatik 2009: 161) could be seen as a kind of parallel to the Russian case, given that in both cases the gender is determined by the location of the geographical feature.

Finally, Enger (2009: 1291f.) highlights that evidence of any parallels observed must be independent. That is to say, typological evidence for a constraint must not be attested only by the same analyst that proposed the constraint. For instance, Steinmetz’s (2006: 1434f.) claim that the principle “functional hollows are neuter” is valid for German due to the candidate nouns being the same gender across many Germanic and Slavic languages is weakened by the fact that the alleged typological parallels are all claimed by him. This would not satisfy Enger’s second criterion as it does not constitute independent evidence.

Method

In order to determine whether there is typological motivation for the constraints which have been proposed for German, the gender assignment literature for three languages with varying degrees of similarity to German, namely Dutch, Norwegian and Spanish, will be studied. It will then be seen whether any gender assignment principles/constraints which are similar to those proposed for German have been established (independently) for these three languages.

It is important to note that there could, of course, be parallel gender assignment principles in languages other than those selected. However, for the purposes of investigating this criterion, studying these three should provide us with an idea of some of the gender assignment principles which have been proposed for other languages, and it will enable us to see whether they are at all similar to those which have been proposed for German.

Results

The typological parallels to constraints invoking semantic features will be dealt with first, before moving on to the parallels found for the constraints pertaining to formal features.

Semantic constraints

The following constraints have parallels in all three of the languages studied, constituting reasonably strong typological evidence for these constraints.

<p>CITIES/TOWNS→*M,*F; COUNTRIES→*M,*F; REGIONS→*M,*F</p> <ul style="list-style-type: none"> • Cities, towns, villages, countries and provinces are neuter in Dutch (Donaldson 2008: 39, Haeseryn et al. 1997: 154). • Cities, towns, villages, countries, provinces, states and regions (apart from those ending in unstressed <i>-a</i>) are masculine in Spanish (Butt & Benjamin 2004: 11). • Nouns denoting geographical administrative units are neuter in Norwegian (Graedler 2013: 218).
<p>MEN→*F,*N; NAME: MAN→*F,*N</p> <ul style="list-style-type: none"> • Nouns denoting men are common gender in Dutch (Donaldson 2008: 34). • Nouns denoting men are masculine in Spanish (Butt & Benjamin 2004: 1). • Nouns denoting men are masculine in Norwegian (Strandskogen & Strandskogen 1986: 57, Trosterud 2001: 36).
<p>WOMEN→*M,*N; NAME: WOMAN→*M,*N</p> <ul style="list-style-type: none"> • Nouns denoting women are common gender in Dutch (Donaldson 2008: 34). • Nouns denoting women are feminine in Spanish (Butt & Benjamin 2004: 1). • Nouns denoting women are feminine in Norwegian (Strandskogen & Strandskogen 1986: 58, Trosterud 2001: 36).
<p>LETTERS→*M,*F</p> <ul style="list-style-type: none"> • Letters of the alphabet are common gender in Dutch (Donaldson 2008: 36, Haeseryn et al. 1997: 153). • Letters of the alphabet are feminine in Spanish (Butt & Benjamin 2004: 10). • Letters of the alphabet are masculine in Norwegian (Trosterud 2001: 40).
<p>TREES→*M,*N</p> <ul style="list-style-type: none"> • Trees are common gender in Dutch (Donaldson 2008: 35, Haeseryn et al. 1997: 153). • Trees (whose fruit is feminine) are masculine in Spanish (Butt & Benjamin 2004: 8). • (Domestic) trees are feminine in Norwegian (Trosterud 2001: 41, Strandskogen & Strandskogen 1986:58).

Parallels to the following constraints were found in Spanish and either Dutch or Norwegian.

<p>DAYS→*F,*N; UNITS OF TIME→*F,*N</p> <ul style="list-style-type: none"> • Days, months and seasons are common gender in Dutch (Donaldson 2008: 35). • Days of the week and months are masculine in Spanish (Butt & Benjamin 2004: 7).
<p>FRUITS→*M,*N; VEG.ROOT→*M,*N</p> <ul style="list-style-type: none"> • Fruits are common gender in Dutch (Donaldson 2008: 35, Haeseryn et al. 1997: 153). • Fruits are feminine in Spanish (Butt & Benjamin 2004: 10).
<p>MUS.NOTES→*M,*F</p> <ul style="list-style-type: none"> • Musical notes are common gender in Dutch (Haeseryn et al. 1997: 153). • Musical notes are masculine in Spanish (Butt & Benjamin 2004: 8).

RIVERS NON-GER.→*F,*N, RIVERS EUROPE→*M,*N; RIVERS NON-EUROPE→*F,*N
<ul style="list-style-type: none"> • Large rivers are common gender in Dutch (Donaldson 2008: 35). • Rivers are masculine in Spanish (Butt & Benjamin 2004: 7). • Towns on the left bank of the Volga are feminine and towns on the right bank are masculine in Russian before 1917 (Superanskaja 1965: 58, cited in Corbett & Fraser 2000: 321).
MALE ANIMALS→*F,*N
<ul style="list-style-type: none"> • Male animals are masculine in Spanish (Butt & Benjamin 2004: 2). • Male animals are masculine in Norwegian (Strandskogen & Strandskogen 1986: 57).

The following constraints have parallels in both Dutch and Norwegian, but not Spanish.

GEN. PERSON→*F,*N
<ul style="list-style-type: none"> • Nouns denoting people (regardless of gender) are common gender in Dutch (Donaldson 2008: 34). • Nouns that refer to a person without specifying their gender are masculine in Norwegian (Trosterud 2001: 36).
LANGUAGES→*M,*F
<ul style="list-style-type: none"> • Languages are neuter in Dutch (Haeseryn et al. 1997: 154). • Languages are neuter in Norwegian (Graedler 2013: 218).
METALS→*M,*F
<ul style="list-style-type: none"> • Metals are neuter in Dutch (Donaldson 2008: 38, Fehringer 1999: 30, Haeseryn et al. 1997: 154). • Materials and substances (including metals) are neuter in Norwegian (Strandskogen & Strandskogen 1986: 58, Trosterud 2001: 40).
MOUNTAINS→*F,*N
<ul style="list-style-type: none"> • Mountains are common gender in Dutch (Donaldson 2008: 35). • Mountains are neuter in Norwegian (Trosterud 2001: 39).
ROCKS/MINERALS→*F,*N
<ul style="list-style-type: none"> • Minerals and stones (where the name denotes the material in general rather than individual stones) are neuter in Dutch (Donaldson 2008: 38). • Rocks and minerals are masculine in Norwegian (Trosterud 2001: 40).
YOUNG ANIMALS→*M,*F
<ul style="list-style-type: none"> • Young animals are neuter in Dutch (Donaldson 2008: 38). • Offspring/young animals are neuter in Norwegian (Strandskogen & Strandskogen 1986: 58).

Parallels to the constraints below were found in just one other language. As can be seen, there is just one constraint which has only a Spanish parallel; all others listed below have a parallel in either Dutch or Norwegian. That is to say, among the constraints relating to semantic properties, where a constraint has a parallel in Spanish, it almost always has a parallel in Dutch and/or Norwegian too.

ISLANDS→*M,*F	Islands are feminine in Spanish (Butt & Benjamin 2004: 10).
ALC. DRINKS→*F,*N	Alcoholic drinks are common gender in Dutch (Franco et al. 2018: 50).
BIRDS→*F,*N	Birds are common gender in Dutch (Donaldson 2008: 35).
FLOWERS→*M,*N	Flowers are common gender in Dutch (Donaldson 2008: 35, Haeseryn et al. 1997: 153).
GAMES→*M,*F	Names of games and sports are neuter in Dutch (Haeseryn et al. 1997: 154; Franco et al. 2018: 50).
MUS. INSTR.→*M,*F; MUS. INSTR.→*M,*N; MUS. INSTR.→*M	Musical instruments are common gender in Dutch (Donaldson 2008: 36, Haeseryn et al. 1997: 153).
SEASONS→*F,*N	Seasons are common gender in Dutch (Donaldson 2008: 35).
CHEMICAL COMPOUNDS/ SUBSTANCES→*M,*F	Names of materials and substances are neuter in Norwegian (Strandskogen & Strandskogen 1986: 58).
DWELL→*M,*F	Permanent residences are masculine in Norwegian (Hjelde 1996: 307f.).
ELEMENTS→*M,*F	Substances (including elements) are neuter in Norwegian (Trosterud 2001: 40).
F. HOLLOWES→*M,*F; OPENINGS→*M,*F	Functional hollows are neuter in Norwegian (Trosterud 2001: 41).

FABRIC→*F,*N LEATHER TYPES→*M,*F	Materials and substances (including fabrics and leather) are neuter in Norwegian (Trosterud 2001: 40).
INSTITUTIONS→*M,*F	Cultural and social institutions are neuter in Norwegian (Trosterud 2001: 39).
ORGANS→*M,*F	Parts of the body are feminine in Norwegian (Strandskogen & Strandskogen 1986: 58).
PLANTS→*F,*N	Plants are masculine in Norwegian (Strandskogen & Strandskogen 1986: 57, Trosterud 2001: 41).
SHEET/STRIP/SLAB→*M,*F	Thin, flat objects/surfaces are neuter in Norwegian (Trosterud 2001: 41).
STICK→*F,*N	Long/oblong objects are masculine in Norwegian (Trosterud 2001: 40).
SUP.→*M,*F	Superordinates are neuter in Norwegian (Trosterud 2001: 37).
WORD CLASSES→*M,*F	Grammatical categories are neuter in Norwegian (Trosterud 2001: 41).

Constraints relating to formal features

“Native” suffixes

The following constraints are those which involve so-called “native” (i.e. Germanic) suffixes. Unsurprisingly, the parallels to these constraints come only from Dutch and Norwegian (and not from Spanish).

-CHEN→*M,*F	<ul style="list-style-type: none"> Nouns with a diminutive suffix are neuter in Dutch (Donaldson 2008: 38, Fehringer 1999: 31).
-ER SUFF.→*F,*N	<ul style="list-style-type: none"> Nouns with the suffix <i>-er</i> are common gender in Dutch (Booij 2002: 38, van der Wouden 2016a). Nouns in <i>-aar</i> are common gender in Dutch (Donaldson 2008: 35). Nouns with the suffix <i>-er</i> are masculine in Norwegian (Askedal 2016: 2537).
GE-→*M,*F	<ul style="list-style-type: none"> Nouns with the prefix <i>ge-</i> are neuter in Dutch (Donaldson 2008: 39, Fehringer 1999: 31, Haeseryn et al. 1997: 154).
-HEIT/-(IG)KEIT→*M,*N	<ul style="list-style-type: none"> Nouns with the suffix <i>-heid</i> are common gender in Dutch (Donaldson 2008: 37, Haeseryn et al. 1997: 152). Nouns with the suffix <i>-het</i> are masculine in Norwegian (Strandskogen & Strandskogen 1986: 57).
-IN SUFF.→*M,*N	<ul style="list-style-type: none"> Nouns with the suffixes <i>-in</i>, <i>-es</i>, <i>-ster</i> etc. are common gender in Dutch (Donaldson 2008: 34f., 54ff.). Nouns with the suffix <i>-inne</i> are feminine in Norwegian (Trosterud 2001: 44).
-LING→*F,*N	<ul style="list-style-type: none"> Nouns in <i>-ling</i> are common gender in Dutch (Donaldson 2008: 36). Nouns with the suffix <i>-ling</i> are masculine in Norwegian (Askedal 2016: 2536).
-NIS→*M,*F; -NIS→*M,*N; -NIS→*M	<ul style="list-style-type: none"> Nouns ending in <i>-nis</i> are common gender in Dutch (Donaldson 2008: 37, Haeseryn et al. 1997: 152).
-SCHAFT→*M,*N	<ul style="list-style-type: none"> Nouns with the suffix <i>-schap</i> signifying a condition are common gender and nouns with the suffix <i>-schap</i> signifying a function are neuter in Dutch (Donaldson 2008: 37, 40). Nouns with the suffix <i>-skap</i> are neuter in Norwegian (Strandskogen & Strandskogen 1986: 58).
-TUM→*F,*N; -TUM→*M,*F; -TUM→*F	<ul style="list-style-type: none"> Nouns with the suffix <i>-dom</i> are neuter in Dutch (Donaldson 2008: 40). Nouns with the suffix <i>-dom</i> are masculine in Norwegian (Strandskogen & Strandskogen 1986: 58).
-UNG→*M,*N	<ul style="list-style-type: none"> Nouns in <i>-ing</i> are common gender in Dutch (Donaldson 2008: 36, Haeseryn et al. 1997: 152).

“Foreign” suffixes

The constraints listed below relate to “foreign” suffixes, i.e. those which have been adopted from other languages, in this case largely from Latin. Parallels to these constraints come from all three of the languages studied.

-ADE→*M,*N	<ul style="list-style-type: none"> Nouns ending in <i>-ade</i> are common gender in Dutch (Haeseryn et al. 1997: 152).
-AGE→*M,*N	<ul style="list-style-type: none"> Nouns in <i>-age</i> are common gender in Dutch (Donaldson 2008: 37, Haeseryn et al. 1997: 152). Nouns in <i>-aje</i> are masculine in Spanish (Butt & Benjamin 2004: 8).
-ANER→*F,*N	<ul style="list-style-type: none"> Nouns with the suffix <i>-(i)aan</i> are common gender in Dutch (van der Wouden 2016b). Nouns with the suffix <i>-aner</i> are masculine in Norwegian (Askedal 2016: 2537).
-ANT→*F,*N	<ul style="list-style-type: none"> Nouns in <i>-ant</i> are common gender in Dutch (Wiktionary 2018a).
-ANZ→*M,*N	<ul style="list-style-type: none"> Nouns in <i>-antie</i> are common gender in Dutch (Wiktionary 2018b).
-AN→*M,*F	<ul style="list-style-type: none"> Nouns in <i>-aan</i> are neuter in Dutch (van der Wouden 2017b).
-AT→*M,*F, -AT→*F,*N	<ul style="list-style-type: none"> Nouns in <i>-aat</i> are neuter in Dutch (Donaldson 2008: 41, Fehring 1999: 31). Nouns in <i>-at</i> are neuter in Norwegian (Trosterud 2001: 48).
-EDER→*M,*F	<ul style="list-style-type: none"> Nouns with the “foreign” suffix <i>-eder</i> are neuter in Norwegian (Wiktionary 2017a).
-EI→*M,*N	<ul style="list-style-type: none"> Nouns in <i>-(er)ij</i> are common gender in Dutch (Donaldson 2008: 37, Haeseryn et al. 1997: 152). Nouns in <i>-eri</i> are neuter in Norwegian (Strandskogen & Strandskogen 1986: 58, Trosterud 2001: 44).
-ENT→*F,*N; -ENT→*M,*F	<ul style="list-style-type: none"> Nouns with suffix <i>-ent</i> are common gender in Dutch (Wiktionary 2018c).
-ENZ→*M,*N	<ul style="list-style-type: none"> Nouns in <i>-entie</i> are common gender in Dutch (Wiktionary 2018b).
-IE [‘i:]→*M,*N	<ul style="list-style-type: none"> Nouns with final stressed <i>-ie</i> [‘i:] are common gender in Dutch (Donaldson 2008: 37, Haeseryn et al. 1997: 152).
-IK→*M,*N	<ul style="list-style-type: none"> Nouns in <i>-iek</i> are common gender in Dutch (Donaldson 2008: 37, Haeseryn et al. 1997: 152).
-INE→*M,*N	<ul style="list-style-type: none"> Nouns ending in <i>-ine</i> are common gender in Dutch (Haeseryn et al. 1997: 152).
-INER→*F,*N	<ul style="list-style-type: none"> Nouns with the suffix <i>-ijn</i> are common gender in Dutch (de Haas & Trommelen 1993: 197f.).
-ING→*M,*F	<ul style="list-style-type: none"> Deverbal (English) nouns in <i>-ing</i> are feminine in Norwegian (Trosterud 2001: 44). Nouns ending in <i>-ing</i> are common gender in Dutch (Franco et al. 2018: 50).
-ION→*M,*N	<ul style="list-style-type: none"> Nouns in <i>-(t)ie</i> are common gender in Dutch (Donaldson 2008: 37). Nouns in <i>-ción, -sión</i> are feminine in Spanish (Butt & Benjamin 2004: 10).
-IST→*F,*N	<ul style="list-style-type: none"> Nouns with the suffix <i>-ist</i> are common gender in Dutch (van der Wouden 2016c).
-ISMUS→*F,*N	<ul style="list-style-type: none"> Nouns ending in <i>-isme</i> are neuter in Dutch (Donaldson 2008: 40).
-ITIS→*M,*N	<ul style="list-style-type: none"> Nouns in <i>-itis</i> are feminine in Spanish (Butt & Benjamin 2004: 10).
-IUM→*M,*F	<ul style="list-style-type: none"> Nouns in <i>-ium</i> are neuter in Dutch (Donaldson 2008: 40). Nouns in <i>-ium</i> are neuter in Norwegian (Strandskogen & Strandskogen 1986: 59).
-IVE→*M,*N	<ul style="list-style-type: none"> Nouns with the suffix <i>-ief</i> are neuter in Dutch, except for linguistic terms in <i>-ief</i> which are common gender (de Haas & Trommelen 1993: 362f.).
-MA→*M,*F	<ul style="list-style-type: none"> Nouns (of Greek origin) ending in <i>-ma</i> are masculine in Spanish (Butt & Benjamin 2004: 9).
-MENT→*M,*F	<ul style="list-style-type: none"> Nouns in <i>-ment</i> are neuter in Dutch (Haeseryn et al. 1997: 153; Franco et al. 2018: 50). Nouns ending in <i>-ment</i> are neuter in Norwegian (Strandskogen & Strandskogen 1986: 59, Graedler 1998: 160, 2013: 218, Trosterud 2001: 44). Nouns ending in <i>-mento</i> and <i>-miento</i> are masculine in Spanish (Tuten et al. 2011: 123).
-(O)LOGE→*F,*N	<ul style="list-style-type: none"> Nouns in <i>-loog</i> are common gender in Dutch (Wiktionary 2017b).
-SKOP→*M,*F	<ul style="list-style-type: none"> Nouns ending in <i>-skop</i> are neuter in Norwegian (Trosterud 2001: 44).
-TÄT→*M,*N	<ul style="list-style-type: none"> Nouns ending in <i>-(i)teit</i> are common gender in Dutch (Donaldson 2008: 37, Haeseryn et al. 1997: 152). Nouns ending in <i>-tet</i> are neuter in Norwegian (Strandskogen & Strandskogen 1986: 59). Nouns ending in <i>-dad</i> are feminine in Spanish (Butt & Benjamin 2004: 10).
-THEK→*M,*N	<ul style="list-style-type: none"> Nouns ending in <i>-theek</i> are common gender in Dutch (Donaldson 2008: 37). Nouns in <i>-teca</i> are feminine in Spanish (Wiktionary 2018d). Nouns in <i>-ek</i> are neuter in Norwegian (Strandskogen & Strandskogen 1986: 59).
-UM→*M,*F	<ul style="list-style-type: none"> Nouns in <i>-um</i> are neuter in Dutch (Donaldson 2008: 40, Fehring 1999: 31, Haeseryn et al. 1997: 153).

	<ul style="list-style-type: none"> Nouns in <i>-um</i> are neuter in Norwegian (Strandskogen & Strandskogen 1986: 59, Trosterud 2001: 48).
-UR SUFF. → *M, *N	<ul style="list-style-type: none"> Nouns in suffix <i>-uur</i> are common gender in Dutch (Wiktionary 2013).

Other formal features

-A → *M, *N	<ul style="list-style-type: none"> Nouns ending in <i>-a</i> are feminine in Spanish (Butt & Benjamin 2004: 2).
-AL → *M, *F	<ul style="list-style-type: none"> Polysyllabic nouns ending in (stressed) <i>-al</i> are neuter in Norwegian (Strandskogen & Strandskogen 1986: 58).
CONVERSION → *M, *F	<ul style="list-style-type: none"> Nominalised infinitives (Donaldson 2008: 38, Fehringer 1999: 31) and words from other categories used as nouns are neuter in Dutch (Haeseryn et al. 1997: 154). All infinitives, quoted words, adverbs, interjections or other genderless words used as a noun are masculine in Spanish (Butt & Benjamin 2004: 7). Quoted words and sound effects are neuter in Norwegian (Trosterud 2001: 400).
DEVERBAL STEM → *F, *N	<ul style="list-style-type: none"> Nouns derived from bare verb stems are common gender in Dutch (Haeseryn et al. 1997: 151). Nouns derived from bare verb stems are neuter in Norwegian (Trosterud 2001: 44).
-E → *M, *N; SUFFIX. -E → *M, *N	<ul style="list-style-type: none"> Nouns ending in <i>-e</i> are feminine in Norwegian (Trosterud 2001: 35). Bisyllabic nouns ending in unstressed <i>-e</i> are feminine in Norwegian (Trosterud 2001: 36).
-FT/-CHT → *M, *N	<ul style="list-style-type: none"> Nouns ending in <i>-ft</i> and <i>-kt</i> are feminine in Norwegian (Trosterud 2001: 47).
-IER /i:r/ POLY. → *M, *F	<ul style="list-style-type: none"> Nouns ending in <i>-ie:r</i> are neuter in Norwegian (Trosterud 2001: 48).
-IV → *F, *N; - IV → *M, *F; -IV → *F	<ul style="list-style-type: none"> Nouns ending in <i>-iv</i> are neuter in Norwegian (Strandskogen & Strandskogen 1986: 59, Trosterud 2001: 48).
-O → *M, *F	<ul style="list-style-type: none"> Nouns ending in <i>-o</i> are masculine in Spanish (Butt & Benjamin 2004: 8).
V _{LONG} C MONO → *F	<ul style="list-style-type: none"> Nouns with a long vowel followed by /d/ are neuter in Norwegian (Trosterud 2001: 48).
V _{STRESSED} [l] → *M, *F	<ul style="list-style-type: none"> Nouns ending in a long high vowel + /l/ are feminine in Norwegian (Trosterud 2001: 47).
V _{STRESSED} [m] → *M, *F	<ul style="list-style-type: none"> Nouns ending in <i>-om</i>, <i>-em</i> and <i>-ym</i> are neuter in Norwegian (Trosterud 2001: 48).
VER → *F; VER- NO SUFF. → *F	<ul style="list-style-type: none"> Nouns formed from verbal stems beginning with the unstressed prefix <i>ver-</i> are neuter in Dutch (Donaldson 2008: 39, Fehringer 1999: 31, Haeseryn et al. 1997: 153).

In total, typological parallels were found in Dutch, Spanish and/or Norwegian for just under half (47%) of the constraints proposed for German. 33% of the constraints had a Dutch equivalent, 26% had a Norwegian equivalent and 12% had an equivalent in Spanish. It is naturally unsurprising that Spanish had the lowest number of equivalent principles, given that it is the most distantly related to German out of the three languages studied.

Of the constraints for which typological evidence was found, 42% had a parallel in at least two of the languages investigated. In all cases, the evidence found was from a source which differed from that in which the principle was originally proposed for German, thereby constituting independent evidence for the constraints.

Almost all of the parallels found were direct parallels to the constraints proposed for German, with only a small number of cases of indirect or near parallels. Aside from the indirect Russian parallel to RIVERS NON-GER. → *F, *N “non-German rivers cannot be feminine or neuter”, there are also cases such as the near Norwegian parallel “parts of the body are feminine in Norwegian” (Strandskogen &

Strandskogen 1986: 58) to the German constraint *ORGANS*→**M*,**F* “names of organs cannot be masculine or feminine” (Steinmetz 1986: 209). Otherwise most parallels were essentially exact.

A surprising finding amongst the results is an independently-attested Norwegian parallel to the constraint *F. HOLLOW*S→**M*,**F* “functional hollows cannot be masculine or feminine”. This is surprising, not only because of the vagueness of the concept upon which the constraint is based, but also because the constraint fared so poorly by the first criterion (with a coverage of just 25%). This highlights a key weakness of the method used to assess constraints by the second criterion, in that the findings are reliant on the assumption that the Dutch, Norwegian and Spanish gender assignment principles cited in the literature are valid. That is to say, the validity of the parallel principles has not been tested; rather, the principles’ citation in the literature has been deemed sufficient evidence for their existence.

When using the literature, including descriptive grammars, to investigate typological parallels, it is important to consider how they might have been influenced. That is to say, it is entirely possible, and indeed quite likely, that many of the ideas found in descriptive grammars (predominantly of European languages) have been influenced not only by those found in other descriptive grammars, but also by the works of early grammarians on classical languages. As stated in Robins (1997), for example, early grammars of European languages were shaped heavily by descriptions of Latin, and it is likely that these ideas have been passed down over time. Therefore, any parallels observed may be a result of grammarians being influenced by each other and/or ideas about classical languages, rather than having independently come to the same conclusions. It is, for instance, interesting that most of the semantic features that have typological support from all three of the languages studied also have a parallel in Latin: nouns denoting women, countries, cities and trees are feminine in Latin, whilst nouns denoting men are masculine (Kennedy 2010: 221).

In sum, it has been established that typological support for constraints can be extremely valuable, and is especially important within standard OT (e.g. Kager 1999: 1, Hayes 2004: 291, McCarthy 2008: 212). However, there are limitations to the methodology used to evaluate constraints by this criterion in that the validity of the principles attested in the three languages studied has not been confirmed, and of course that the gender assignment principles of only three languages were investigated. Therefore, the results may be used as an indication of the existence of typological parallels, but not as irrefutable proof. In any case, Enger (2009: 1291) states that fulfilment of this criterion is not essential for a constraint, since many will be without typological parallels, and ‘one cannot *a priori* exclude language-particular generalisations’.

2.5 Criterion 3: Parallels outside of gender assignment

‘If the rule invokes a particular “semantic feature”, it is an advantage if we recognise this feature from other parts of the grammar of the language in question’ (Enger 2009: 1292)

With his third criterion, Enger (2009) seeks evidence that the feature upon which a constraint is based is relevant to German outside of gender assignment. Presumably the logic behind this criterion is that if a feature is relevant to one area of the language, then we know that the feature is already recognised and used by native speakers. Therefore, the use of the feature within gender assignment is made more plausible. This idea is also touched on by Mills (1986: 114), who considers the ‘clarity’ of a gender assignment rule to be improved if the principle is related to ‘additional functions’ (i.e. outside of gender assignment), and Corbett (1991: 31), who observes that ‘several of the criteria which underlie gender systems also turn up regularly in other aspects of morphology and syntax’.

Enger (2009: 1292) states that this criterion is part of the reason why constraints based on features such as animacy seem more plausible than those based on, for instance, dairy products or stick-like objects. However, he acknowledges that many valid principles will fare badly by this criterion and therefore that fulfilment of this criterion is not essential, since ‘the possibility of rules based on “unique” features cannot be ruled out *a priori*’ (Enger 2009: 1292).

One example of a constraint which involves a feature that is relevant to German outside of gender assignment is CITIES/TOWNS → *M, *F “the names of cities/towns cannot be masculine or feminine”, e.g. *das junge Berlin* ‘the young Berlin’ or *das schöne Paris* ‘the beautiful Paris’ (e.g. Eisenberg 2013: 138, Köpcke & Zubin 2005: 119, Weinrich 2007: 328). In addition to being relevant to gender assignment, city/town names are also relevant to adjectival inflection. When used as attributive adjectives, they do not inflect like regular adjectives; they always take an *-er* ending, regardless of the gender, case or number of the noun they are modifying, e.g. *der Berliner Platz* ‘the Berlin square’, *die Berliner Mauer* ‘the Berlin wall’, *mit dem Berliner Ensemble* ‘with the Berlin ensemble’ (e.g. Duden Grammatik 2009: 753). Accordingly, this constraint seems to satisfy criterion three.

Although Enger (2009) proposes this criterion for constraints based on semantic features only, it can also be extended to formal constraints, since there are also formal features that are relevant to German outside of gender assignment. The criterion can therefore also partially account for why formal constraints based on features such as suffixes are more readily acceptable than those based on, for instance, a seemingly arbitrary cluster of phonemes.

Method

To assess the constraints using this criterion, key words or segments from the constraints will be searched for in the digital edition of the Duden Grammatik (2009) to determine whether the features they involve are relevant to any area of the language other than gender (e.g. phonology, morphology or syntax). Accordingly, any gender feature – formal or semantic – which is potentially related to any feature in German other than gender is listed below. As with Section 2.4, the constraints relating to semantic features will be dealt with first, followed by the constraints involving formal features.

Results

Constraints involving semantic features

Noun formation

The following constraints all involve semantic features which are relevant to the formation of nouns.

WOMEN→*M,*N <ul style="list-style-type: none">• The suffix <i>-in</i> may be added to many nouns denoting people to signify that the referent is a woman, e.g. <i>Lehrerin</i> ‘teacher [woman]’ (Duden Grammatik 2009: 732).• With compounds ending in <i>-mann</i>, the element <i>-mann</i> is often replaced by <i>-frau</i> if the referent is a woman, e.g. <i>Fachfrau</i> ‘expert [woman]’ (Duden Grammatik 2009: 156).
MEN→*F,*N <ul style="list-style-type: none">• For compounds ending in <i>-mann</i>, e.g. <i>Geschäftsmann</i> ‘businessman’, the plural form should be <i>-männer</i> if the referents are all men. <i>-leute</i> is used to refer to a mixed-gender group (Duden Grammatik 2009: 156).
INSTITUTIONS→*M,*F <ul style="list-style-type: none">• The (foreign) suffix <i>-(i)at</i> derives the names of institutions from nouns denoting people, e.g. <i>Kommissariat</i> ‘commissariat’, <i>Dekanat</i> ‘deanery’ (Duden Grammatik 2009: 731).
MALE ANIMALS→*F,*N <ul style="list-style-type: none">• The suffix <i>-(e)rich</i> may be added to nouns denoting an animal in order to derive the term for the male, e.g. <i>Enterich</i> ‘drake’ (Duden Grammatik 2009: 732).
NAME: MAN→*F,*N; NAME: WOMAN→*M,*N <ul style="list-style-type: none">• Unisegmental shortened words usually consist of the initial segment of the full form. With personal names, however, the shortered form usually consists of the final segment, e.g. <i>Tina</i> < <i>Christina</i>, <i>Bert</i> < <i>Robert</i> (Duden Grammatik 2009: 735).

Adjective formation

The constraints below invoke features which are relevant to the formation of adjectives.

FABRIC→*F,*N <ul style="list-style-type: none">• For the formation of adjectives meaning ‘made of (a certain material)’, the suffix <i>-(e)n/-ern</i> can be attached to certain nouns. One group of nouns to which this suffix can be added are the names of fabrics, e.g. <i>seiden</i> ‘silk’, <i>samten</i> ‘velvet’ (Duden Grammatik 2009: 758).
UNITS OF TIME→*F,*N <ul style="list-style-type: none">• Adjectives can be formed from units of time plus <i>-lich</i> (and sometimes an umlauted vowel), e.g. <i>monatlich</i> ‘monthly’, <i>jährlich</i> ‘yearly’, <i>stündlich</i> ‘hourly’. Such adjectives cannot be used predicatively (Duden Grammatik 2009: 359).
TIMES OF DAY→*F,*N <ul style="list-style-type: none">• Adjectives can be formed from nouns denoting a time of day plus <i>-lich</i>, e.g. <i>abendlich</i> ‘in the evenings’ (Duden Grammatik 2009: 756).

ROCKS/MINERALS→*F,*N

- For the formation of adjectives meaning 'made of (a certain material)', the suffix *-(e)n/-ern* can be attached to certain nouns. One group of nouns to which this suffix can be added are the names of minerals, e.g. *diamanten* 'diamond', *golden* (Duden Grammatik 2009: 758).

CITIES/TOWNS→*M,*F; REGIONS→*M,*F; COUNTRIES→*M,*F

- Adjectives derived from place names (mostly cities, sometimes regions or countries) are formed with suffix *-er*. These adjectives are only used attributively and do not inflect like regular adjectives, in that the suffix *-er* is used irrespective of the gender, case and number of the noun it is modifying, e.g. *die Londoner U-Bahn* 'the London underground', *in dem Hamburger Hauptbahnhof* 'in Hamburg's central station'.
- The suffix *-(i)sch* derives adjectives from the names of most countries, e.g. *mazedonisch* 'Macedonian' (Duden Grammatik 2009: 753).

Agreement

The following constraints contain features which are relevant to agreement.

GEN. PERSON→*F,*N

- Among nouns referring to people, there are often variants for referring to men (e.g. *Besitzer* 'owner [man]') and variants for referring to women (e.g. *Besitzerin* 'owner [woman]'). Where this is the case, in order to refer to a generic person, the variant for referring to men should be used, e.g. *Jedes Mitglied ist Besitzer eines Vereinsausweises* 'every member is the owner of a membership card' (Duden Grammatik 2009: 996).

GROUPS→*M,*N

- When a collective term (e.g. *Gruppe* 'group', *Menge* 'crowd', *Bande* 'gang' etc.) in the singular has a plural attribute, the finite verb may be singular (syntactic agreement) or plural (semantic agreement). The plural is more common with collective terms which have a 'vague' meaning. The singular is more common when the attribute is not mentioned (Duden Grammatik 2009: 1014).

Formation of temporal adverbs

TIMES OF DAY→*F,*N; DAYS→*F,*N

- Temporal adverbs can be formed from a noun denoting a time of day or a noun denoting a day of the week plus *-s*, e.g. *abends* 'in the evenings', *montags* 'on Mondays' (Duden Grammatik 2009: 575).

Noun countability and plural formation

MON. UNITS→*F,*N

- For most monetary units (often those which are masculine or neuter), no plural ending is required, e.g. *zehn Euro* 'ten euros', *hundert Dollar* 'a hundred dollars'. Plural endings may be used in more concrete contexts, e.g. *ich tausche meine Dollars gegen Euros* 'I'm exchanging my dollars for euros'.
- The plural ending *-en* is always used for monetary units to which it applies. These are often feminine.
- For foreign monetary units, the plural form from the original language is sometimes used (Duden Grammatik 2009: 177f.).

SCI. UNITS→*M,*F

- Most scientific units appear in the singular, even when they are used in a plural sense, e.g. *hundert Gramm* (**Gramme*) 'a hundred grams'.
- Accordingly, the dative plural *-n* is not added to scientific units when in the singular, even if they are being used in a plural sense, e.g. *mit drei Gramm* (**Grammen*) 'with three grams'. This, however, does not apply to scientific units ending in unstressed *-el* or *-er*, which may take the dative plural *-n*, since these can also be interpreted as plural forms, e.g. *mit drei Liter(n)* 'with three litres' (Duden Grammatik 2009: 177f.).

LETTERS→*M,*F

- Individual letters take an optional *-s* plural in German, e.g. *ein Wort mit drei F(s)* 'a word with three Fs' (Duden Grammatik 2009: 189).

NAME: MAN→*F,*N; NAME: WOMAN→*M,*N

- When forming the plural of personal names, it is possible to use the rules for common nouns or to use an *-s* plural, e.g. *die Adelheiden/die Adelheids*, *die Fritzchen/die Fritzchens* (Duden Grammatik 2009: 191)

Use of articles

CITIES/TOWNS→*M,*F; REGIONS→*M,*F; COUNTRIES→*M,*F

- Articles are not generally used with the names of regions or countries, e.g. *ich fahre nach Spanien* 'I'm going to Spain'. However, there are a number of exceptions to this, including when placenames are feminine or plural, e.g. *die Lausitz* 'Lusatia', *die Niederlande* 'the Netherlands' (Duden Grammatik 2009: 299f.).

NAME: MAN→*F,*N; NAME: WOMAN→*M,*N

- Articles are not generally used with personal names, e.g. *Hallo Stefan!* However, there are a number of exceptions to this, including when names appear with modifiers, e.g. *die kluge Anna* 'the clever Anna', or to express distance or derogation, e.g. *Goethe oder nur der Schiller?* 'Goethe or just Schiller?' (Duden Grammatik 2009: 301f.).

Nominalised adjectives and case marking

LANGUAGES→*M,*F

- Names of languages can appear either in the 'nominalised form', e.g. *das Deutsch* 'German' or the 'adjectival form', e.g. *das Deutsche* 'German'. Their distribution is determined by a number of semantic and formal factors. The adjectival form can only be used when weakly inflected (i.e. with *-e* or *-en* endings). In instances where strong inflection would be required, the nominalised form is used, e.g. *Er versteht das Deutsche* 'he understands German', but *er versteht kein Deutsch* 'he doesn't understand German' (Duden Grammatik 2009: 205, 350).
- In the genitive, names of languages in the nominalised form take either the short *-s* ending or no ending at all, e.g. *des Deutsch(s)*. However, it is more common for the adjectival form to appear after a definite article, demonstrative or possessive determiner, e.g. *die Verbesserung Ihres Englischen* 'the improvement of her English' (rather than *Ihres Englisch(s)*) (Duden Grammatik 2009: 205, 350).

LETTERS→*M,*F

- In the genitive, individual letters take either the short *-s* ending or no ending at all, e.g. *des langen i/des langen Is* 'of the long i' (Duden Grammatik 2009: 205).

CITIES/TOWNS→*M,*F; REGIONS→*M,*F; COUNTRIES→*M,*F

- With placenames, case endings appear only in the genitive. The dative and accusative forms have no ending. In the genitive, only the short ending *-s* is found, e.g. *Berlins bekanntestes Hotel* 'Berlin's most famous hotel', *der höchste Berg Australiens* 'the highest mountain in Australia'. Most place names are not subject to the genitive rule stating that a noun may only appear in the genitive if it is preceded by an inflected article or adjective (Duden Grammatik 2009: 208).

Formation of demonyms

CITIES/TOWNS→*M,*F; REGIONS→*M,*F; COUNTRIES→*M,*F

- Demonyms for cities are usually formed with suffix *-er*, e.g. *Leipziger*. There are a few cities for which this is not always the case, e.g. *Tokioter* 'Tokyoite'.
- The demonym for countries ending in *-land* or *-en* is usually formed by adding *-e*, e.g. *Schwede* 'Swede'. For countries ending in an open syllable, the demonym is formed by adding *-aner*, e.g. *Mexikaner* 'Mexican'. The demonym for countries whose last consonant is a sonorant is formed by adding *-ese*, e.g. *Vietnamese*. For countries ending in *-ien*, the demonym is formed with the suffixes *-e*, *-er* or *-aner*, e.g. *Kroate* 'Croatian', *Spanier* 'Spaniard', *Kolumbianer* 'Colombian' (Duden Grammatik 2009: 729f.).

Prepositional adverbs

GROUPS→*M,*N

- Prepositional adverbs are generally only used when they are referring to an inanimate. Should an animate be the referent, a preposition together with a personal pronoun must be used instead, e.g. *Ich warte auf den Auftrag* 'I'm waiting for the order' → *Ich warte darauf (*auf ihn)* 'I'm waiting for it' vs. *Ich warte auf meinen Hund* 'I'm waiting for my dog' → *Ich warte auf ihn (*darauf)* 'I'm waiting for him'. However, if a group of people (or animals) is being referred to, a prepositional adverb may also be used: *Es kam eine Gruppe von Wanderern*; *Otto steuerte direkt darauf/auf sie zu* 'a group of hikers came; Otto headed straight for them' (Duden Grammatik 2009: 582).

Constraints relating to formal features

Noun formation

The constraints below all involve formal features which are all relevant to the formation of nouns in some way.

The following constraints all involve “native” derivational affixes which can be added to a base to create nouns (Duden Grammatik 2009: 723).

-CHEN→*M,*F	GE-→*M,*F	-NIS→*M,*N	-SCHAFT→*M,*N
-EL→*F,*N	-HEIT/-(IG)KEIT→*M,*N	-NIS→*M	-TUM→*F,*N
-EL→*M,*F	-IN SUFF.→*M,*N	-SAL→*M	-TUM→*M,*F
-EL→*M,*N	-LING→*F,*N	-SAL→*M,*F	-TUM→*F
-ER SUFF.→*F,*N	-NIS→*M,*F	-SAL→*M,*N	-UNG→*M,*N

The following constraints all contain “foreign” derivational suffixes which can also be added to a base to create nouns (Duden Grammatik 2009: 723).

-ADE→*M,*N	-EE/-ÉE/-É→*M,*F	-EUR/ÖR→*N	-ISMUS→*F,*N
-AGE→*M,*N	-EE/-ÉE/-É→*M,*N	-ETTE→*M,*N	-MENT→*M,*F
-ANT→*F,*N	-EE/-ÉE/-É→*F,*N	-IE [‘i:’]→*M,*N	-(O)LOGE→*F,*N
-ANZ→*M,*N	-EI→*M,*N	-IK→*M,*N	-OR→*F,*N
-AR SUFF.→*F,*N	-ENZ→*M,*N	-INE→*M,*N	-TÄT→*M,*N
-AR SUFF.→*M,*F	-ENT→*F,*N	-ING→*M,*F	-UM→*M,*F
-AT→*M,*F	-ENT→*M,*F	-ION→*M,*N	-UR SUFF.→*M,*N
-AT→*F,*N	-EUR/ÖR→*F,*N	-IST→*F,*N	

The following constraints involve “commercial” suffixes which are found in various product names (Duden Grammatik 2009: 741).

-ON→*M,*F	e.g. <i>Amiodaron, Risperidon, Valoron</i>
-ON UNSTRESS.→*M,*F	
-IN [i:n]→*M,*F	e.g. <i>Aspirin, L-Thyroxin, Alpecin</i>

The following constraints are also relevant to noun formation.

-SKOP→*M,*F	- <i>skop</i> is a combining form (a bound morpheme with lexical meaning) which is used to create nouns (Duden Grammatik 2009: 682).
-THEK→*M,*N	- <i>thek</i> is a combining element used to form nouns. It is usually preceded by the linking element - <i>o-</i> (Duden Grammatik 2009: 682-685).
CONVERSION→*M,*F	Conversion is a common method for noun formation (Duden Grammatik 2009: 725).
DEVERBAL STEM→*F,*N	Creating nouns from verb stems is also a common method for noun formation, e.g. <i>schlafen</i> ‘to sleep’ > <i>der Schlaf</i> ‘the sleep’ (Duden Grammatik 2009: 667).
DEVERBAL -T→*M,*N	Noun formation using verb stems plus - <i>t</i> , e.g. <i>fahren</i> ‘to travel’ > <i>die Fahrt</i> ‘journey’, is a process which has been used to create many German nouns (Duden Grammatik 2009: 678).

Word stress

The constraints below all contain suffixes which affect word stress since they are stress-attracting suffixes (Duden Grammatik 2009: 49).

-AGE→*M,*N	-AT→*F,*N	-ION→*M,*N	-IV→*F
-ANT→*F,*N	-EI→*M,*N	-IST→*F,*N	-IV STRESSED→*M,*F
-ANZ→*M,*N	-ENZ→*M,*N	-IV→*F,*N	-TÄT→*M,*N
-AT→*M,*F	-INE→*M,*N	-IV→*M,*F	-UR SUFF.→*M,*N

Genitive inflection

The following constraints all involve formal features which are relevant to genitive inflection. The set listed below usually take the “short” genitive ending *-s* (Duden Grammatik 2009: 299).

-IG→*F,*N; -ING→*M,*F; -LING→*F,*N	Nouns ending in <i>-ig</i> , <i>-ing</i> or <i>-ling</i> , e.g. <i>des Königs</i> , <i>des Campings</i> , <i>des Flüchtlings</i> .
-A→*M,*N; -O→*M,*F	Nouns ending in a full vowel, e.g. <i>des Klimas</i> , <i>des Kontos</i> .
-EL→*F,*N etc.; -EN→*F,*N etc.; -ER→*F,*N etc.	Nouns ending in unstressed <i>-el</i> , <i>-en</i> or <i>-er</i> , e.g. <i>des Segels</i> , <i>des Balkens</i> , <i>des Lehrers</i> .

The following kinds of noun usually take the “long” genitive ending *-es* (Duden Grammatik 2009: 97, 196f.).

-ST/-SCHT/-ZT→*N; -ST/-SCHT/-ZT→*F,*N	Nouns ending in <i>-st</i> (but the short ending is also permissible), e.g. <i>des Diensts/des Dienstes</i> .
-C S MONO→*F,*N	Nouns ending in [s], e.g. <i>des Kreuzes</i> , <i>des Witzes</i> , <i>des Fuchses</i> .
-(C)(C)/J/ MONO→*F,*N	Nouns ending in <i>-sch</i> , e.g. <i>des Busches</i> (the short ending is also permissible).
MONOSYLL.→*F,*N	Monosyllables are more likely to display the “long” genitive ending <i>-es</i> than, for example, bisyllabic words which have stress on the first syllable, e.g. <i>des Triebes</i> (92%) – <i>des Triebs</i> (8%) vs. <i>des Antriebes</i> (20%) – <i>des Antriebs</i> (80%).

Compound formation

Where the first element in a compound ends in one of the following suffixes, the linking element *-s* normally follows, e.g. *Schönheit-s-wettbewerb*, *Diversität-s-index*, *Museum-s-neubau* (Duden Grammatik 2009: 651, 714f., 751).

-HEIT/-(IG)KEIT→*M,*N	-TÄT→*M,*N	-TUM→*F
-LING→*F,*N	-TUM→*F,*N	-UM→*M,*F
-SCHAFT→*M,*N	-TUM→*M,*F	

Plural formation

Nouns with the following features generally take no ending in the plural (Duden Grammatik 2009: 181f., 186).

-CHEN→*M,*F	Nouns with suffix <i>-chen</i> , e.g. <i>das Blümchen</i> > <i>die Blümchen</i> .
-EN→*F,*N etc. -ER→*F,*N etc. -EL→*F,*N etc.	Nouns ending in unstressed <i>-el</i> , <i>-en</i> or <i>-er</i> , e.g. <i>der Schatten</i> > <i>die Schatten</i> .
GE→*M,*F	Nouns ending in unstressed <i>-e</i> usually take the plural ending <i>-en</i> . However, nouns in <i>-e</i> which are prefixed with <i>Ge-</i> do not follow this rule. Instead, they take no plural ending, e.g. <i>das Gewebe</i> > <i>die Gewebe</i> .

The following kinds of nouns take an optional *-s* plural (Duden Grammatik 2009: 189)

CONVERSION→*M,*F	Converted nouns, e.g. <i>viele Wenn(s) und Aber(s)</i> .
------------------	--

The constraints below contain a feature which is relevant to the *-e* plural (Duden Grammatik 2009: 184).

-NIS→*M,*F etc. -SAL→*M etc.	Feminine nouns with an <i>-e</i> plural or a plural with no ending are umlauted (where possible), e.g. <i>die Wand</i> > <i>die Wände</i> . Exceptions to this rule are nouns with the suffixes <i>-sal</i> and <i>-nis</i> , which take an <i>-e</i> plural without an umlaut, e.g. <i>die Drangsal</i> – <i>die Drangsale</i> .
---------------------------------	---

The following constraints contain features which are relevant to the *-en* plural form (Duden Grammatik 2009: 182, 186f.).

-UM→*M,*F	Nouns ending in <i>-um</i> usually replace the <i>-um</i> with <i>-en</i> in the plural, e.g. <i>Museum</i> > <i>Museen</i> .
-ON UNSTRESS.→*M,*F; -ON→*M,*F	Nouns ending in <i>-on</i> usually replace the <i>-on</i> with <i>-en</i> in the plural, e.g. <i>Stadion</i> > <i>Stadien</i> .
-IS→*M,*N	Nouns ending in <i>-is</i> usually replace the <i>-is</i> with <i>-en</i> in the plural, e.g. <i>Praxis</i> > <i>Praxen</i> .
-E→*M,*N; SUFFIX.- E→*M,*N	Nouns ending in <i>-e</i> usually take the plural suffix <i>-en</i> , e.g. <i>Seite</i> > <i>Seiten</i> .
-(O)LOGE→*F,*N	Nouns in <i>-(o)loge</i> take an <i>-n</i> in the plural, e.g. <i>Biologe</i> > <i>Biologen</i> .
-A→*M,*N; -O→*M,*F	Usually, nouns whose singular form ends in an unstressed full vowel take an <i>-s</i> in the plural. Exceptions to this rule are often nouns of foreign origin ending in <i>-a</i> and <i>-o</i> , which take an <i>-en</i> ending, e.g. <i>Firma</i> > <i>Firmen</i> , <i>Risikos</i> > <i>Risiken</i> (<i>Risikos</i> is also possible but less common).
-IUM→*M,*F	Nouns in <i>-ium</i> usually end in <i>-ien</i> in the plural, e.g. <i>Kolloquium</i> > <i>Kolloquien</i> .
-IE [i:]→*M,*N; -IE [iə]→*M,*N	Nouns ending in <i>-ie</i> take <i>-n</i> in the plural, e.g. <i>Galerie</i> > <i>Galerien</i> , <i>Studie</i> > <i>Studien</i> .
-US→*F,*N	Nouns ending in <i>-us</i> often take a replaceive <i>-en</i> in the plural, e.g. <i>Rhythmus</i> > <i>Rhythmen</i> .

Dative plural inflection

The constraints below contain formal features which are relevant to dative plural inflection.

-EL→*F,*N etc.; -ER→*F,*N etc.	<ul style="list-style-type: none"> The dative plural <i>-n</i> cannot be added to singular nouns. However, (singular) scientific units ending in unstressed <i>-el</i> or <i>-er</i> may take the dative plural <i>-n</i>, since these can also be interpreted as plural forms, e.g. <i>mit drei Liter(n)</i> 'with three litres' (Duden Grammatik 2009: 176).
--------------------------------	---

Demonym formation

-ER SUFF. → *F, *N; -ANER → *F, *N

- *-er* and *-aner* are suffixes which can be used to derive demonyms from countries ending in *-ien*, e.g. *Albaner* ‘Albanian’, *Brasilianer* ‘Brazilian’ (Duden Grammatik 2009: 730).

Weak masculine nouns

The following constraints involve properties which are relevant to weak masculine nouns.

-E → *M, *N; SUFFIX.-E → *M, *N

- Weak masculine nouns are often those which denote an animate and end in *-e* in the nominative singular (Duden Grammatik 2009: 212f.).

The suffixes below are commonly found among weak masculine nouns (Duden Grammatik 2009: 213).

-IST → *F, *N	e.g. <i>der Tourist</i>
-IT → *F, *N; -IT → *M, *F; -IT → *F	e.g. <i>der Bandit</i>
-ANT → *F, *N	e.g. <i>der Demonstrant</i>
-AT → *M, *F; -AT → *F, *N	e.g. <i>der Kandidat</i>
-(O)LOGE → *F, *N	e.g. <i>der Biologe</i>

To summarise, of all 94 meaning-based constraints relevant to the noun sample, the features of 20 were found to be relevant to areas of the language other than gender. The semantic features of these constraints are mostly relevant to areas such as the formation of adjectives and nouns, plural inflection, genitive inflection and noun countability. Of the 143 form-based constraints relevant to the noun sample, the features of 79 were found to be relevant to other areas of the language. These areas include: noun formation, genitive and plural inflection and compound formation.

2.6 Criterion 4: Clarity of semantic features

‘If the rule invokes a particular “semantic feature”, it is an advantage if we recognise this feature from elsewhere, and it should be reasonably clear what nouns the rule covers’ (Enger 2009: 1292)

Enger’s (2009: 1292) reasoning behind the fourth criterion is that: ‘a rule that relies on a feature we have come across before, if only outside of the grammar, is more plausible than one which relies on a feature we never have come across before.’ That is to say, gender assignment principles which draw on culturally-relevant categories are more likely to be recognised and therefore used by native speakers than those which rely on vague concepts which are difficult to distinguish. Enger (2009: 1292) states that this criterion explains why principles based on semantic categories such as “alcoholic beverages”, for instance, seem more acceptable than those based on e.g. “functional hollows”.

The requirement for constraints to be clear is also advocated by Comrie (1999: 461), who states that ‘in some cases the principles that are adduced are of such complexity that it is questionable whether they are valid’. He reasons that principles which are based on highly complex and/or vague features are unlikely to be valid, since it is doubtful that such principles would be able to be extracted by native speakers in the acquisition process. Fulfilment of this criterion thus seems essential for semantic constraints, given that if it is not clear to which nouns a constraint is relevant, it is unlikely that the constraint would be picked up on and used by native speakers.

As with Enger’s (2009) other criteria, the main challenge with criterion four is how to practically determine whether or not a constraint satisfies it. In terms of devising a method for this purpose, the criterion can be divided into two parts. The first is assessing whether the semantic feature upon which a constraint is based can be recognised from elsewhere, and the second is assessing whether it is “reasonably clear” to which nouns a constraint is relevant. In terms of the first part, Enger (2009: 1292) suggests that if a feature is recognisable then there will be extralinguistic evidence for it. Whilst, as Enger (2009: 1292) states, ‘we admittedly do not possess any list of what is a culturally relevant category’, sources such as encyclopaedias and internet search engines can help to determine whether or not there is extralinguistic evidence for a category. The second part, on the other hand, is related to the notion of the clarity of category membership, a theoretical matter which will be considered in the second half of this section.

Method

In order to investigate the first part of criterion four, extralinguistic evidence for the semantic features referred to by each constraint must be sought in order to determine whether the features are culturally relevant and thus recognisable to the speech community. This will be done by using sources such as the Oxford-Duden pictorial dictionary (1994), online encyclopaedias (e.g. Wikipedia, Wissen.de etc.) and a variety of other websites (listed in full in Appendix D) to see whether or not they make reference to the features upon which the constraints are based.

Results

For the majority of constraints involving a semantic feature (72 out of 94), extralinguistic evidence was found to suggest that the categories they invoke are of cultural relevance and therefore recognisable by the speech community. Most of the evidence was found easily using the aforementioned sources, since many of the categories are well established, e.g. birds, trees, musical instruments, countries etc. A few of the categories proved slightly more difficult to find evidence for, such as pictorial images (PICTURES→*M,*F) and stick-like objects (STICK→*F,*N), perhaps because these categories are

somewhat less obvious, but there was nonetheless extralinguistic evidence to be found for all of these. See Appendix D for the full list of evidence for these constraints.

Extra-linguistic evidence could not, however, be found for the following 22 semantic categories, listed below alongside examples of candidate nouns cited by the constraints' proponents. See Appendix B for full source information.

Constraint	Semantic category
ANNOUNCE→*M,*F	Announcements/pronouncements of an official or public nature, e.g. <i>das Bulletin</i> , <i>Verdikt</i> 'verdict'.
CONDUCT→*M,*N	Abstract principles as guides to conduct, e.g. <i>die Doktrin</i> 'doctrine', <i>Regel</i> 'rule'.
ELAB. INSTR.→*M,*N	Instrumental objects with a functionally elaborated end, e.g. furcated or pincer-like objects, e.g. <i>die Gabel</i> 'fork', <i>Zange</i> 'pliers'.
EXTROVERSION→*F,*N	Nouns denoting extroversion, e.g. <i>der Zorn</i> 'anger', <i>Mut</i> 'courage', <i>Eifer</i> 'zeal'.
F. HOLLOW→*M,*F	Functional hollows, e.g. <i>das Auge</i> 'eye', <i>Rad</i> 'wheel', <i>Tor</i> 'gate'.
FEAR→*M,*N	Nouns denoting fear, anxiety or urgency, e.g. <i>die Angst</i> 'fear', <i>Not</i> 'emergency'.
IMPROPER SUPERORDINATES→*M,*F	Improper superordinates, which 'represent a purely mental classification' requiring 'judgements of a human mind', e.g. <i>das Ziel</i> 'aim', <i>Rätsel</i> 'riddle', <i>Wunder</i> 'miracle'.
INTEGRATED PARTS→*F,*N	Integrated parts, e.g. <i>der Bestandteil</i> 'component'.
INTROVERSION→*M,*N	Nouns denoting introversion, e.g. <i>die Scheu</i> 'shy', <i>Geduld</i> 'patience', <i>Anmut</i> 'grace'.
LIVING/WORKING→*M,*F	Living and working spaces, e.g. <i>das Haus</i> 'house', <i>Quartier</i> 'quarters'.
MESH→*M,*F	Mesh-like structures or cruciform objects, e.g. <i>das Netz</i> 'net', <i>Sieb</i> 'sieve'.
NO INFO.→*M,*F	Nouns which provide no specific information about their referent, e.g. <i>das Ding</i> 'thing', <i>Objekt</i> 'object', <i>Zeug</i> 'stuff'.
OPEN WATER→*M,*N	Open bodies of water, e.g. <i>die See</i> 'sea'.
OPENINGS→*M,*F	Openings, e.g. <i>das Becken</i> 'basin', <i>Fenster</i> 'window', <i>Grab</i> 'grave'.
PCC GARMENT→*M,*F	Primary chest-covering garments: the garment must cover the area between the waist and the chest and must not be outerwear, e.g. <i>das Kleid</i> 'dress', <i>Hemd</i> 'shirt'.
POWER/STRENGTH→*M,*N	Nouns denoting power and strength, e.g. <i>die Gewalt</i> 'force'.
SEMANTIC DIMINUTIVES→*M,*N	Semantic diminutives, referring to 'smaller, weaker, less active and less prominent' entities, e.g. <i>die Insel</i> 'island' (vs. <i>das Land</i> 'country'), <i>die Tür</i> 'door' (vs. <i>das Tor</i> 'gate'), <i>die Matte</i> 'mat' (vs. <i>der Teppich</i> 'carpet').
SEPARABLE PARTS→*M,*F	Nouns denoting separable parts, e.g. <i>das Einzelteil</i> 'part'.
SHEET/STRIP/SLAB→*M,*F	Sheets, strips or slabs, e.g. <i>das Blatt</i> 'leaf', <i>Blech</i> 'sheet', <i>Brett</i> 'board'.
SPEECH→*M,*N	Acts of speech, e.g. <i>die Frage</i> 'question', <i>Rede</i> 'speech'.
SUPERORDINATES OF INDIRECT REFERENCE→*M,*F	Superordinates of indirect reference, defined by 'vagueness of the denotata'. The category is said to 'reflect a basic property of neuter nouns in general (highlighted by Zubin & Köpcke 1984: 144) that neuter nouns index lesser differentiation in their referents and categorize objects at the highest level of generalisation', e.g. <i>das Erlebnis</i> 'experience', <i>Risiko</i> 'risk', <i>Bedürfnis</i> 'need', <i>Desaster</i> 'disaster'.
TIME: SHORT→*F,*N	Short periods of time, e.g. <i>der Tag</i> 'day'.

For a few of the constraints, e.g. SEMANTIC DIMINUTIVES→*M,*N, the lack of extralinguistic evidence is most likely due to the fact that the features are defined in relative terms, e.g. "smaller entities" is not objectively verifiable. Some arguably borderline cases among the results are "nouns denoting extroversion" (EXTROVERSION→*F,*N), "introversion" (INTROVERSION→*M,*N) and, to some extent, "fear, anxiety and urgency" (FEAR→*M,*N), since it is possible to find synonyms for each of these terms. However, there is no extralinguistic evidence for these terms serving as categories, and no evidence could be found which groups together the concepts of fear, anxiety and urgency.

Since there is no extralinguistic evidence for the categories referred to by these constraints, this suggests that they are not recognisable by the speech community, and it is consequently unlikely that the constraints are able to be acquired or used by native speakers. These constraints therefore fail to satisfy the first part of criterion four.

Whether or not categories are recognisable by the speech community is also related to the second part of criterion four, namely how clear it is which nouns the categories cover. That is to say, part of the reason why a category might not be recognised by a speech community could be that the category is unclear. Equally, a category might seem unclear to an individual principally because that particular category is not recognised by their speech community.

It is somewhat difficult to test the second part of the criterion, since whether or not it is “reasonably clear” which nouns are covered by a category is a subjective matter. Nevertheless, some general problems that exist in relation to the clarity of category membership will be discussed below using examples of the constraints being tested.

The main issue concerning the clarity of category membership is the method of determining whether a noun is a member of a category or not. The traditional approach to categorisation is based on the Aristotelian idea of necessary and sufficient conditions (e.g. Gärdenfors 1999: 25, Taylor 2003: 20ff.). According to this approach, categories are defined by a set of necessary, binary conditions which together are sufficient for defining the category (e.g. Riemer 2010: 224f.). In order to be considered a member of a particular category, an entity must satisfy all of these conditions. As a result, category membership is seen as a binary issue, which means that categories have clear boundaries (e.g. Löbner 2002: 175, Taylor 2003: 21). According to this model, it should therefore be extremely clear which nouns are covered by a category if a set of necessary and sufficient conditions can be established.

For some of the categories under scrutiny here, this approach may seem satisfactory. For instance, the category WOMEN could be defined by the necessary and sufficient conditions of being human, adult and identifying as a woman. However, for other categories this model proves inadequate. For example, the category BIRDS might reasonably be defined by a set of conditions such as: has wings, flies, sings, has feathers and lays eggs (e.g. Riemer 2010: 224). Yet these conditions cannot be necessary or sufficient, since, for example, many birds do not fly (ostrich, emu, penguin etc.) and not all birds have songs (e.g. gulls, ducks) (e.g. Taylor 2003: 38). Moreover, what if there were, for example, a robin born without wings due to a genetic mutation? Does it then cease to belong to the category BIRDS? (See Kripke (1980: 119ff.) for a similar example about tigers).

In addition, it has been argued that there are some categories which cannot be defined by any necessary conditions at all, since there is no single feature that is shared by all members. For example, Wittgenstein (1953) claims that there are no properties which are common to all members of the category GAMES. Instead, some games share features with certain others. For instance, some games require physical activity, some involve chance, some are played strategically and some involve a board, but, importantly, none of these features are true for all games, yet all games are related. Wittgenstein (1953) describes the relation between all members of the category GAMES as a 'family resemblance', since, like a family, they are alike as a group but no one feature is common to all members.

Since determining sets of necessary and sufficient conditions often proves problematic, it seems reasonable to conclude that the Aristotelian approach to categorisation is not suitable for all categories. An alternative approach to categorisation which appears to overcome many of the problems of necessary and sufficient conditions is Prototype Theory, as developed primarily by Rosch (1973, 1975a, 1975b), Rosch & Mervis (1975), Rosch et al. (1976), Mervis & Rosch (1981) etc.⁶

According to Prototype Theory, categories are structured around a central prototype, which is considered to be the "best example" of a category (Rosch 1973, 1975a, 1975b). Accordingly, categories are claimed to have a graded structure, which allows for the existence of more and less typical examples of a particular category (e.g. Rosch & Mervis 1975, Rosch 1998). This is supported by experimental evidence from Rosch (1975b: 229-232), who asked 200 individuals to rate items in terms of how good an example of a category they were. The results showed that there was a high level of agreement across the group. For example, in the category FRUITS, an orange was generally considered to be a better example than an avocado, among BIRDS, a robin was considered a more typical example than a penguin, and football was considered to be a better example than billiards in the category SPORTS.

For some categories, it therefore seems sensible to discuss category membership in terms of "normally sufficient conditions" or "typical conditions" rather than as a fixed set of necessary and sufficient conditions (e.g. Cruse 2000: 56). If, for instance, the conditions "has wings, flies, sings, has feathers and lays eggs" are seen as typical rather than necessary for membership of the category BIRDS, it prevents non-prototypical birds from being excluded from the category.

With regard to the categories invoked by the constraint set, it seems that it is the clarity of these conditions, along with the ease at which they can be established, that determines how clear it is which

⁶ For a more comprehensive discussion of Prototype Theory and subsequent developments of the theory, see e.g. Evans & Green (2006: 248-282).

nouns belong to the category. For instance, it could be argued that the reason ALCOHOLIC DRINKS might be seen as a clearer category than SHEETS, STRIPS AND SLABS is because the typical conditions for determining membership of the former are clearer than those for determining membership of the latter. For ALCOHOLIC DRINKS, one might reasonably assume the typical conditions “is a drink, contains alcohol”. For SHEETS, STRIPS AND SLABS, however, the conditions are unclear. Must the item consist of a single sheet, strip or slab, for instance? Or would an item which is comprised of a collection of sheets, strips or slabs, e.g. a book or a garden fence, be considered a member of the category?

It is important to note, however, that for some categories, the typically sufficient conditions might well be clear, but they may not be explicitly known to all members of the speech community. For example, most would consider the category METALS to be relatively clear in terms of the nouns that it covers. It is not difficult to name some substances which are metals and some which are not. However, it is likely that most people are in fact unaware of the exact scientific conditions which typically define membership of the category METALS. Despite this lack of knowledge, we are still able to employ this category in a meaningful way. This is because we rely on the fact that there are scientific experts in the speech community who are able to specify the typical conditions of the category. This phenomenon is known as “the division of linguistic labour” (Putnam 1975). It is therefore important to take into consideration that the conditions which typically define a category can be clear, even if they are not known to every individual in the speech community. This is likely to be the case for categories in the constraint set such as ELEMENTS, METALS, ROCKS AND MINERALS, TREES etc.

Related to this point is a further issue concerning the clarity of category membership, namely whether categorisation is indeed based on scientific classification or whether it is instead based on the cultural perception of items, which may not always align with science. For instance, culturally, tomatoes are generally considered to be part of the category VEGETABLES, despite scientifically belonging to FRUIT. Similarly, whales are often thought of as belonging to FISH rather than MAMMALS. Defining categories in terms of prototypical conditions means that such borderline cases could be considered non-prototypical members of either category. For this reason, categories such as FRUIT and MAMMALS may be considered to be slightly less clear than others.

To summarise, fulfilment of Enger’s (2009) fourth criterion is important for constraints invoking semantic features. If a semantic feature is unrecognisable and unclear, then it is unlikely to be picked up on or used by native speakers for gender assignment. In terms of practically assessing constraints by this criterion, the first part, namely recognisability by native speakers, can be determined more easily than the second, namely clarity of category membership. As suggested by Enger (2009: 1292),

recognisability of a feature by native speakers is indicated by the presence of extralinguistic evidence for that feature. All but 22 semantic constraints satisfied this part of the criterion. Clarity of category membership is more difficult to establish for each constraint given the subjectivity of the issue. While general problems concerning category membership have been identified, the clarity of individual categories cannot be objectively quantified. Therefore, fulfilment of criterion four will be judged according to satisfaction of the first part of the criterion.

2.7 Criterion 5: Accounting for exceptions

‘If rule A is necessary in order to account for exceptions from a well-established rule B, this is indirect evidence for A’ (Enger 2009: 1293)

Enger’s fifth criterion is based on a claim made by Zubin & Köpcke (1984: 45) that exceptions to gender assignment principles are ‘the consequence of competition with other [...] principles’ (an idea which is also central to OGAT). Enger (2009: 1293) reasons that, as a consequence of this, exceptions constitute indirect evidence for the other principles. For example, he states that the masculine exceptions to the Norwegian principle “nouns ending in unstressed /e/ are feminine”, *rømme* ‘sour cream’, *dravle* ‘curd’ and *myse* ‘whey’, provide indirect evidence for the principle “dairy products are masculine” (Trosterud 2001, Nessel 2006a), since they are instances of the semantic principle overriding the formal principle.

An equivalent German example might be the neuter exceptions to the phonological principle “monosyllables ending in the consonant clusters /tʃ/, /Ntʃ/ or /Nj/ (whereby *N* signifies a nasal) are masculine” (Salmons 1993: 425), namely *das Deutsch* ‘German’ and *das Romansch* ‘Romansh’, which provide indirect evidence for the principle “names of languages are neuter” (e.g. Flämig 1991: 453, Köpcke & Zubin 1996: 480, Menzel 2004: 63 etc.), since they are the result of the latter principle overriding the former.

Method

To test the constraints against the fifth criterion, it is necessary to look for cases of constraints accounting for exceptions to other constraints. In order to do this, the exceptions identified when investigating the first criterion (Section 2.3) will be analysed, and exceptions whose gender might be explained by the presence of other constraints will be highlighted as indirect evidence for these other constraints.

Enger (2009: 1293) states that there is only indirect evidence for a principle if it accounts for exceptions from another “well-established” principle. While he does not specify what constitutes a “well-established” principle, in this case we are able to use any constraint which satisfies Yang’s (2005) Tolerance Principle (that is, passes criterion one), since for these constraints, their exceptions really are exceptions, i.e. a minority of cases. Therefore, only exceptions to constraints which pass the first criterion will be examined.

Of the constraints found in 2.3 to satisfy the Tolerance Principle, 117 have at least one exception. The exceptions to these 117 constraints (3,343 nouns in total) have then been analysed to see if they can be accounted for by any other constraint(s) in the set.

Results

It was found that approximately two thirds of the exceptions could be accounted for by other constraints in the set, thus providing indirect evidence for these other constraints. In total, evidence for 101 constraints was found.

It should, of course, be noted that, since there are GENDER FEATURES constraints which are not being evaluated in this chapter as they are not relevant to the noun sample (see 2.2), it is possible that the remaining third of exceptions could be accounted for by these untested constraints. Similarly, it is possible that there is further indirect evidence for some of the constraints tested in the form of exceptions to these untested GENDER FEATURES constraints.

The table below shows the constraints which are able to account for a substantial number of exceptions, indicating that there is relatively strong indirect evidence for these constraints. The first column shows the constraints for which there is indirect evidence according to the fifth criterion, the second column shows the total number of exceptions accounted for by this constraint, and the third column shows some examples of the exceptions accounted for, including the constraint to which the exceptions belong and the exceptional nouns themselves.

Constraint for which there is indirect evidence	No. of exceptions accounted for	Examples
GEN. PERSON → *F, *N	423	-AT → *M, *F: <i>der Delegat, Diplomat, Stipendiat</i> -E → *M, *N: <i>der Biologe, Serbe, Zeuge</i> -IK → *M, *N: <i>der Katholik, Domestik</i>
GE- → *M, *F	144	-E → *M, *N: <i>das Gebilde, Gebirge</i> -ER → *F, *N: <i>das Gejammer, Geschwister</i>
-(O)LOGE → *F, *N	84	-E → *M, *N: <i>der Biologe, Musikologe</i>
-FT/-CHT → *M, *N	71	TR- → *F, *N: <i>die Tracht, Trift</i>
-HEIT/-IG)KEIT → *M, *N	66	GE- → *M, *F: <i>die Geeignetheit, Gefährlichkeit</i>

LANGUAGES→*M,*F	58	-E→*M,*N: <i>das Baskische, Finnische</i> -(C)(C)/J/ MONO→*F,*N: <i>das Deutsch</i>
DEVERBAL -T→*M,*N	54	CCVCC→*F,*N: <i>die Flucht, Schlacht, Schrift</i> DEVERBAL STEM→*F,*N: <i>die Schrift, Sicht, Ankunft</i>
-E→*M,*N	34	SUP.→*M,*F: <i>die Waffe, Pflanze, Farbe</i> FABRIC→*F,*N: <i>die Seide, Viskose, Wolle, Chenille</i>
MONOSYLL.→*F,*N	24	GAMES→*M,*F: <i>der Skat, Jass</i>
CVCC MONO→*F		-ST/-SCHT/-ZT→*N: <i>das Nest, Fest, Biest</i>
CONVERSION→*M,*F	23	-E→*M,*N: <i>das Zuhause, Heute</i>
-A→*M,*N	21	MON. UNITS→*F,*N: <i>die Griwna, Kuna, Lira, Rufiyaa, Stotinka</i> RIVERS NON-GER.→*F,*N: <i>die Angara, Wolga, Moskwa</i>
-IN SUFF.→*M,*N	20	GE→*M,*F: <i>die Gefährtin, Gehilfin</i>
CVC MONO→*F	19	V _{LONG} C MONO→*F: <i>die Kur, Schur</i>
-UNG→*M,*N		GE→*M,*F: <i>die Genesung, Gewährung</i>
MEN→*F,*N	15	-E→*M,*N: <i>der Junge, Neffe</i>
SUP.→*M,*F	14	-ER→*F,*N: <i>das Geschwister, Wetter, Desaster</i> -EKT→*F,*N: <i>das Insekt, Konfekt</i>
CCVC MONO→*F,*N		GAMES→*M,*F: <i>der Skat</i>
ALC. DRINKS→*F,*N		-E→*M,*N: <i>der Sake, Amarone, Cidre</i>
-ER SUFF.→*F,*N		GE→*M,*F: <i>der Gesellschafter, Gewinner</i> GAMES→*M,*F: <i>der Bettler, Kicker</i>
ROCKS/MINERALS→*F,*N		-IN [i:n]→*M,*F: <i>der Olivin, Zitrin, Erythrin</i>
NAME: MAN→*F,*N	10	-E→*M,*N: <i>Schiele, Rilke, Mörrike, Macke</i> -A→*M,*N: <i>Kalidasa, Garcia Lorca</i>

It is unsurprising that most of these constraints appear in the above table, since almost all of these were found to have a relatively high percentage coverage of candidate nouns in Section 2.3.

The constraints in the table below are able to account for slightly fewer exceptions, namely between 2 and 9 exceptions each, showing that there is at least some indirect evidence for these constraints according to the fifth criterion. Constraints accounting for just one exception have not been recorded, since this does not constitute substantial indirect evidence for these constraints, and also does not meet Enger's (2009: 1293) requirement that a constraint must account for 'exceptions' in the plural.

Constraint for which there is indirect evidence	No. of exceptions accounted for	Examples
FABRIC→*F,*N	9	-ETT→*M,*F: <i>der Mokett</i>
-EL→*F,*N	8	ELEMENTS→*M,*F: <i>der Schwefel</i>
ELEMENTS→*M,*F		-N(C)(C) MONO→*F,*N: <i>das Zink, Brom, Zinn</i>
SCI. UNITS→*M,*F		-N(C)(C) MONO→*F,*N: <i>das Phon, Ohm, Pfund</i>
-EN→*M,*F; -EN→*F		FABRIC→*F,*N: <i>das Leinen</i>
CHEMICAL COMP./SUBST.→*M,*F		-N(C)(C) MONO→*F,*N: <i>das Zink, Brom, Zinn</i>
DOM. ANIMALS→*M,*F		-N(C)(C) MONO→*F,*N: <i>das Schwein, Huhn, Lamm</i>
-ER→*F,*N	7	TREES→*M,*N: <i>der Wacholder</i>
V _{LONG} MONO→*F		COUNTRIES→*M,*F: <i>der Tschad</i>
V _{LONG} C MONO→*F		GROUPS→*M,*N: <i>der Clan/Klan, der Chor</i>
MESH→*M,*F		-C S MONO→*F,*N: <i>das Kreuz, Netz</i>
METALS→*M,*F		-[I]C MONO→*F,*N: <i>das Gold</i>

GAMES→*M,*F	6	-ETTE→*M,*N: <i>das Roulette</i>
DWELL→*M,*F		CCVCC→*F,*N: <i>das Zelt</i>
GROUPS→*M,*N		CCVCC→*F,*N: <i>die Plebs</i>
CHEESES→*F,*N		-ETTE→*M,*N: <i>der Mimolette</i>
-CHEN→*M,*F		MEN→*F,*N: <i>das Kerlchen, Männchen</i>
-EI→*M,*N		COUNTRIES→*M,*F: <i>die Slowakei, Türkei</i>
-SCHAFT→*M,*N		GE→*M,*F: <i>die Gewerkschaft, Gesellschaft</i>
INSTITUTIONS→*M,*F		5
CCVV _{DIPHTHONG} - MONO→*F	[[]]C MONO→*F,*N: <i>das Schwein</i>	
DEVERBAL STEM→*F,*N	GROUPS→*M,*N: <i>der Verband</i>	
NAME: WOMAN→*M,*N	VCC MONO→*F,*N: <i>Els</i>	
NO INFO.→*M,*F	-N(c)(c) MONO→*F,*N: <i>das Ding, Dings</i>	
CITIES/TOWNS→*M,*F	-E→*M,*N: <i>(das) Skopje</i>	
MAMMALS→*F,*N	-E→*M,*N: <i>der Bulle, der Löwe</i>	
MUS. TERMS→*M,*F	-E→*M,*N: <i>das Graduale, Timbre, Andante</i>	
-IS→*M,*N	REGIONS→*M,*F: <i>die Kolchis, Nearktis</i>	
FLOWERS→*M,*N	4	
FEAR→*M,*N		V _{LONG} C MONO→*F: <i>die Not</i>
[[]]C - MONO→*F,*N		NO INFO.→*M,*F: <i>der Stück</i>
CVCC MONO→*N		SPEECH→*M,*N: <i>der Satz</i>
-EL→*M,*N	3	GEN. PERSON→*F,*N: <i>die Geisel</i>
-EN→*F,*N; -EN→*F		GE→*M,*F: <i>der Gedanken</i>
F. HOLLOWES→*M,*F		-C S MONO→*F,*N: <i>das Herz</i>
DISEASES→*M,*N		R- MONO→*F: <i>die Ruhr</i>
YOUNG ANIMALS→*M,*F		-[[]]C MONO→*F,*N: <i>das Kalb</i>
INSECTS→*M,*N		MALE ANIMALS→*F,*N: <i>die Drohne</i>
-N(c)(c) MONO→*F,*N		GROUPS→*M,*N: <i>der Clan/Klan</i>
-L MONO→*F		MESH→*M,*F: <i>der Grill, Tüll</i>
-ANG→*F,*N		GE→*F,*N: <i>der Gesang</i>
VCCC MONO→*M		-N(c)(c) MONO→*F,*N: <i>die Angst</i>
-IE [i:]→*M,*N		FABRIC→*F,*N: <i>die Ramie</i>
-UM→*M,*F		GEN. PERSON→*F,*N: <i>das Individuum</i>
CONDUCT→*M,*N		-IN [i:n]→*M,*F: <i>die Doktrin, Disziplin</i>
LIVING/WORKING→*M,*F		-N(c)(c) MONO→*F,*N: <i>das Amt</i>
BIRDS→*F,*N		-IN [i:n]→*M,*F: <i>der Pinguin</i>
PLANTS→*F,*N		-IN [i:n]→*M,*F: <i>der Jasmin, Rosmarin</i>
PRIMATES→*F,*N		-E→*M,*N: <i>der Affe, der Schimpanse</i>
WOMEN→*M,*N		-ER→*F,*N: <i>die Mutter, Tochter</i>
-IT→*M,*F; -IT→*F		ROCKS/MINERALS→*F,*N: <i>das Bromit, Wulfenit</i>
COUNTRIES→*M,*F		-E→*M,*N: <i>(das) Zimbabwe, Zaire, Belize</i>
∅V _{LONG} MONO→*F	VC MONO→*M,*F: <i>der Aal, Aar, Öhm</i>	
-LING→*F,*N	2	FRUITS→*M,*N: <i>der Speierling</i>
-AT→*M,*F		GROUPS→*M,*N: <i>das Dezernat</i>
PICTURES→*M,*F		-IK→*M,*N: <i>das Mosaik</i>
-C S MONO→*F,*N		SPEECH→*M,*N: <i>der Satz</i>
R- MONO→*F		CONDUCT→*M,*N: <i>der Rang</i>
T- MONO→*F		MESH→*M,*F: <i>der Tüll</i>
V _{SHORT} STOP C→*F,*N		SPEECH→*M,*N: <i>der Satz</i>
SUFFIX -E→*M,*N		GE→*M,*F: <i>die Gebärde</i>
VV _{DIPHTH} C MONO→*F		CONDUCT→*M,*N: <i>der Brauch</i>
-/r/ STOP (C) MONO→*F,*N		INTROVERSION→*M,*N: <i>der Schmerz</i>
VC MONO→*M,*F		SCI. UNITS→*M,*F: <i>das Ohm</i>
DR→*F,*N		CONDUCT→*M,*N: <i>der Drill</i>
-ÜR/-ÜHR /y:r/→*M,*N		T- MONO→*F: <i>die Tür</i>
-EHR/-ÄHR→*M,*N		GE→*M,*F: <i>die Gewähr</i>
ACAD. DISCIPLINES→*M,*N		-N(C)(C) MONO→*F,*N: <i>die Kunst</i>

-ILLE→*M,*N		FABRIC→*F,*N: <i>die Chenille</i>
MUS. COMP.→*M,*F		-E→*M,*N: <i>das Andante, Graduale</i>
LEATHER TYPES→*M,*F		-ER→*F,*N: <i>das Juchtenleder, Kalbleder</i>
FISH→*F,*N		-E→*M,*N: <i>der Pelamide, Elritze</i>
-EL→*M,*F		GEN. PERSON→*F,*N: <i>das Mündel, Ekel</i>
-IV→*M,*F; -IV→*F; -IV		GROUPS→*M,*N: <i>das Kollektiv, Kooperativ</i>
STRESSED→*M,*F		
-ETT→*M,*F		GROUPS→*M,*N: <i>das Oktett, Terzett</i>

Again, it is largely unsurprising that most of these constraints are supported by the fifth criterion, since the majority also have a high percentage coverage of candidate nouns (Section 2.3). However, there are several for which this is not true, including: INSTITUTIONS→*M,*F (38%); F. HOLLOWES→*M,*F (25%); and -EL→*M,*F (24%). This demonstrates that satisfying the fifth criterion alone is not sufficient evidence for the validity of a constraint, given that it could be used to support unlikely constraints such as F. HOLLOWES→*M,*F, which has thus far failed criteria 1, 3 and 4.

Enger's (2009: 1293) assertion that accounting for exceptions constitutes only 'indirect evidence' for a principle is therefore key. The fifth criterion should accordingly not be seen as a test of the validity of a constraint, but rather a means of providing further confirmation for constraints which have been shown to be valid elsewhere. This is particularly important, given that accounting for exceptions is most likely to be the very reason why certain dubious principles were proposed in the first instance. For example, it is likely that F. HOLLOWES→*M,*F was created to explain nouns such as *das Auge* 'eye' (an exception to -E→*M,*N), *das Rad* 'wheel' and *das Ohr* 'ear' (exceptions to MONOSYLL.→*F,*N). Therefore, criterion five is most useful when used in conjunction with Enger's (2009) other criteria. The relative importance of each of the criteria will be explored further in 2.9.

In sum, the fifth criterion can be a useful measure of constraint plausibility, since if all nouns with a certain feature belong to one gender, except for those which also have another feature, then this indirectly highlights the relevance of the other feature to gender assignment. In total, indirect evidence for 101 constraints was found. However, it was determined that the fifth criterion is best used alongside other criteria because accounting for exceptions alone is not sufficient evidence for a GENDER FEATURES constraint.

2.8 Criterion 6: Productivity

'If a rule is diachronically productive (i.e. if new nouns are assigned the gender predicted by a particular "crazy rule", or if old nouns change gender in the direction that a particular crazy rule predicts), this is an indication that this rule is valid' (Enger 2009: 1293)

The sixth and final criterion suggests that a gender assignment principle is more likely to be valid if it is productive. Productivity is a notoriously problematic term, since, as stated by e.g. Mayerthaler (1981: 124), it is one of the most unclear concepts in linguistics. However, in this instance, Enger (2009: 1293) states clearly that a principle is productive either if new nouns are assigned the gender predicted by the principle, or if there is evidence of nouns having changed gender in the direction that a principle would predict.

Enger (2009: 1293) claims that, for example, the Italian principle “nouns denoting cars are feminine” (Thornton 2009: 25) satisfies the first part of this criterion, since new nouns of this type are indeed feminine in Italian. He also claims that the Norwegian principle “nouns denoting trees are feminine” (Trosterud 2001) satisfies the second part of this criterion, since there is evidence, provided by Beito (1954: 39), that nouns such as *alm* ‘elm’, *ask* ‘ash’, *hassel* ‘hazel’ and *rogn* ‘rowan’ are among the nouns which have changed from masculine to feminine in many Norwegian dialects.

It is clear that evidence of the productivity of a gender assignment principle is extremely valuable, since this is an indication that the principle is psychologically real and synchronically active in the gender assignment system of native speakers. Heringer (1995: 213), too, highlights the importance of productivity, stating that if a gender assignment principle is applicable to new nouns, this means that the principle is valid.

However, as with Enger’s (2009) other criteria, no specific methods have been suggested for assessing constraints by the sixth criterion. It is, for example, not stated how many or what proportion of neologisms must follow a principle in order for the principle to be classed as productive. Similarly, it is also not specified how many nouns must have changed gender in the direction dictated by a principle for the principle to be classed as productive. In terms of overall numbers, we can, however, reason that there must be evidence of at least more than one noun conforming to a principle in order to suggest that a pattern is occurring. In terms of calculating what might constitute a sufficient proportion of nouns, it is once again appropriate to use Yang’s (2005) Tolerance Principle, since this has been designed for estimating the proportion of lexical items that a productive principle should cover. Specific methods for each part of the criterion will be developed in 2.8.1 and 2.8.2.

In addition, it is also not stated by Enger (2009) whether productivity ought to be dealt with here as a binary feature, i.e. a principle is either productive or unproductive (as advocated by e.g. Booij (1977: 5) and Zwanenburg (1983: 29)), as a matter of degree on a continuum (e.g. Bauer 1992), or as on a scale with the discrete values “productive”, “semi-productive” and “unproductive” (e.g. Matthews 1974: 52, 222; Pinker & Prince 1991: 231). However, for this particular criterion, it is appropriate to

treat productivity as a binary feature, since, ideally, it needs to be determined whether there is sufficient evidence to suggest that a gender assignment principle is productive or not. However, there may well be borderline cases, which are of course better dealt with using a continuum approach.

Section 2.8.1 will assess the constraints according to the first part of the criterion, i.e. productivity as indicated by the gender of neologisms, and Section 2.8.2 will assess the constraints according to the second part of the criterion, namely productivity as indicated by historical gender changes. Note that the productivity of constraints will be tested further in Chapter 4 via an experiment involving pseudo-nouns.

2.8.1 Neologisms

In order to determine whether the constraints fulfil the first part of the criterion, a corpus of neologisms has been used to ascertain whether or not there is evidence of new nouns conforming to the constraints. The corpus used is an online dictionary of neologisms created by the Institut für Deutsche Sprache, accessed via the OWID portal (Online-Wortschatz-Informationssystem Deutsch). The neologism dictionary currently lists 1,811 words which have entered the German language since 1990. 1,466 of these 1,811 lexical items are nouns. The table below shows the distribution of genders across all nouns in the corpus.

Gender	No. of nouns ⁷	Proportion of whole corpus
Masculine	596	37.7%
Feminine	443	28.1%
Neuter	540	34.2%

The corpus was then searched for neologisms with features invoked by each of the constraints being assessed. The proportion of neologisms with the relevant feature that have been assigned the gender predicted by the constraint was then calculated. It should be noted that compound neologisms whose head is not a neologism itself were discounted. Given that German compounds assume the gender of the head noun (see 1.3.1.4 and 3.3.3), compounds whose heads are not new nouns themselves cannot be considered new evidence for a constraint. For instance, the neologism *das Boxspringbett* has not been classed as evidence of a new noun following the principle $-ETT \rightarrow *M, *F$, since its gender is determined by the established head noun *das Bett* 'bed'.

⁷ The figures are calculated such that multiple gender nouns are counted for each gender they can be used with, e.g. *der/das Rave* is counted twice; once as masculine and once as neuter. Of the 1,466 nouns in the neologism dictionary, 119 are listed as multiple gender nouns.

In total, evidence for 64 constraints was found. For the constraints listed in the table below, all corresponding neologisms were found to have the gender predicted by the principle, strongly indicating that these principles are productive.

Constraint	No. of neologisms in OWID	Percentage with predicted gender	Satisfies Tolerance Principle?	Examples
-ER SUFF. → *F, *N	134	100%	✓	<i>der Geocacher, der Twitterer</i>
CONVERSION → *M, *F	105	100%	✓	<i>das Kitesurfen, das Dissen</i>
-ING → *M, *F	84	100%	✓	<i>das Streaming, das Spinning</i>
-IN SUFF. → *M, *N	37	100%	✓	<i>die Boulderin, die Bloggerin</i>
-UNG → *M, *N	11	100%	✓	<i>die Aldisierung, die McDonaldisierung</i>
-EI → *M, *N	8	100%	✓	<i>die Youtuberei, die Twitterei</i>
-MENT → *M, *F	7	100%	✓	<i>das Attachment, das Outplacement</i>
-ION → *M, *N	7	100%	✓	<i>die Location, die Preselection</i>
-LING → *F, *N	6	100%	✓	<i>der Ichling, der Schübling</i>
-CHEN → *M, *F	6	100%	✓	<i>das Vürstchen</i>
-HEIT/-(IG)KEIT → *M, *N	6	100%	✓	<i>die Prolligkeit</i>
-TÄT → *M, *N	5	100%	✓	<i>die Metrosexualität</i>
-IE ['i:] → *M, *N	3	100%	✓	<i>die Orthorexie</i>
-OR → *F, *N	2	100%	✓	<i>der Rollator</i>

The above findings are fairly unsurprising, since all of these principles were found to have a 100% coverage of candidate nouns in 2.3. There are, however, a number of principles which could not account for the genders of all candidate nouns in Section 2.3 yet also show signs of being productive. These are shown in the table below. In all cases, the proportion of neologisms with the gender predicted by the principle is higher than or equal to the proportion required by the Tolerance Principle.

Constraint	No. of neologisms in OWID	Neologisms with predicted gender	Satisfies Tolerance Principle?	Examples
-EN → *F	17	100%	✓	<i>das Kitesurfen</i>
-TUM → *M, *F; -TUM → *F	12	100%	✓	<i>das Adrenalinjunkietum, das Nerdium</i>
STICK → *F, *N	5	100%	✓	<i>der Bengalo, der Pen</i>
-IK → *M, *N	4	100%	✓	<i>die Mechatronik</i>
CVCC MONO → *F	4	100%	✓	<i>der Nerd</i>
T- MONO → *F	4	100%	✓	<i>das Taupe</i>
-B → *F	4	100%	✓	<i>das Web</i>
MEN → *F, *N	4	100%	✓	<i>der Mr. Right, der Anchorman</i>
WOMEN → *M, *N	3	100%	✓	<i>die DJane, die Anchorwoman</i>
MON. UNITS → *F, *N	2	100%	✓	<i>der Euro</i>
DAYS → *F, *N	2	100%	✓	<i>der Girls' Day, der Boys' Day</i>
FRUITS → *M, *N	2	100%	✓	<i>die Cranberry</i>
-AL → *M, *F	2	100%	✓	<i>das Tribal</i>
GE- → *M, *F	2	100%	✓	<i>das Gegoogel</i>
-IS → *M, *N	2	100%	✓	<i>die Aufschieberitis</i>
-AR → *M, *F	2	100%	✓	<i>das Webinar</i>
GEN. PERSON → *F, *N	129	98%	✓	<i>der Digital Native, der Mingle</i>

GAMES→*M,*F	36	89%	✓	<i>das Tae-Bo, das Pilates</i>
V _{LONG} MONO→*F	8	88%	✓	<i>der Stream, das Zine</i>
CVC MONO→*F	22	86%	✓	<i>der Pen, das Net</i>
-N(C)(C) MONO→*F,*N	7	86%	✓	<i>der Pen, der Stream</i>
V _{LONG} C MONO→*F	7	86%	✓	<i>der Nerd, das Nude, das Taupe</i>
CCVV _{DIPHTHONG} - MONO→*F	5	80%	✓	<i>der Blade, der Skate, das Trike</i>
-AT→*M,*F	4	75%	✓	<i>das Prekariat, das Plastinat</i>
-ON→*M,*F	4	75%	✓	<i>das Macaron, das Emoticon</i>
GROUPS→*M,*N	4	75%	✓	<i>die Girlgroup</i>
VV _{DIPHTH} C MONO→*F	15	73%	✓	<i>der Kite, das Gate, das Trike</i>
-EN→*M,*F	17	71%	✓	<i>das Kitesurfen</i>
PICTURES→*M,*F	9	56%	✓	<i>das Emoticon</i>
DWELL→*M,*F	4	50%	✓	<i>das Townhouse, das Floating Home</i>
-L MONO→*F	4	50%	✓	<i>der Proll</i>
V _{SHORT} STOP C→*F,*N	2	50%	✓	<i>der Ex</i>
-/r/STOP (C) MONO→*F,*N	2	50%	✓	<i>der Nerd</i>
WASTE→*F,*N	2	50%	✓	<i>der Spam</i>
TR→*F,*N	2	50%	✓	<i>der Trash</i>
CONDUCT→*M,*N	2	50%	✓	<i>die Political Correctness</i>
ALC. DRINKS→*F,*N	2	50%	✓	<i>der Hugo</i>
CCCVC MONO→*F,*N	2	50%	✓	<i>der Stream</i>

For several of the constraints, the findings are complicated somewhat by the presence of multiple gender nouns (MGNs, see 1.1). As is expected, a number of the neologisms in the corpus are listed as being able to be used with more than one gender. For the principles in the table below, the percentage of neologisms listed with only the gender predicted by the principle is lower than the percentage required by the Tolerance Principle. However, if MGNs are included in the calculations such that a neologism is deemed to follow a principle if it lists the predicted gender as one of its possible genders, the percentages increase to a level greater than or equal to the percentages required by the Tolerance Principle, meaning that these principles could be considered to be productive. MGNs are discussed further in 3.3.4.8.

Constraint	No. of neologisms in OWID	Neol. with only predicted gender	Neol. with predicted gender incl. MGNS	Tolerance Principle requirement		Examples
MUSIC TYPES→*F,*N	2	0%	100%	0% ⁸	✓	<i>der/das Techno</i>
CCVC MONO→*F,*N	18	50%	78%	72%	✓	<i>der Croc, der/das Blog</i>
CVCC MONO→*N	4	50%	75%	50%	✓	<i>der Nerd, der/das Link</i>
SUP.→*M,*F	8	50%	75%	63%	✓	<i>der/das Event</i>
-EL→*M,*N	14	50%	64%	64%	✓	<i>das/die Sequel</i>
LIVING/WORKING→*M,*F	7	43%	57%	57%	✓	<i>das Homeoffice</i>
-E→*M,*N	7	43%	57%	57%	✓	<i>die Gabione</i>
DEVERBAL STEM→*F,*N	18	33%	83%	67%	✓	<i>der Chat, der/das Download</i>
SHEET/STRIP/SLAB→*M,*F	13	31%	62%	62%	✓	<i>das Whiteboard, das Post-it</i>

⁸ When the total number of nouns is very low, i.e. 1 or 2, the Tolerance Principle technically allows all nouns to be exceptions. However, in this case it is clear that the constraint is not productive unless MGNS are considered, since there are no neologisms which have only the gender predicted by the constraint.

PREP. DISHES→*M,*F	4	25%	50%	50%	✓	<i>der/das Wrap</i>
-(C)(C)// MONO→*F,*N	4	25%	50%	50%	✓	<i>der/das Slush</i>

Finally, there are a number of constraints which appear to be unproductive based on the evidence available from neologisms. For all of the constraints listed below, the percentage of neologisms with the gender predicted by the principle, even when MGNs are included in the calculations, is considerably lower than the percentage required by the Tolerance Principle. These constraints therefore fail the first part of the sixth criterion.

Constraint	No. of neologisms in OWID	Neologisms with predicted gender	Neologisms with predicted gender incl. MGNs	Percentage required by Tolerance Principle	Satisfies Tolerance principle?
-A→*M,*N	13	31%	46%	62%	x
-EN→*F,*N	17	29%	29%	65%	x
-AT→*F,*N	4	25%	25%	50%	x
PCC GARMENT→*M,*F	4	25%	25%	50%	x
DISEASES→*M,*N	9	22%	22%	56%	x
-EL→*F,*N	14	21%	29%	64%	x
-ER→*M,*N	19	16%	16%	68%	x
-EL→*M,*F	14	14%	21%	64%	x
F. HOLLOWES→*M,*F	16	13%	31%	69%	x
OPENINGS→*M,*F	16	13%	31%	69%	x
VER-→*F	16	6%	6%	69%	x
-TUM→*F,*N	12	0%	0%	67%	x
MESH→*M,*F	2	0%	0%	0% ⁹	x
PLANTS→*F,*N	2	0%	0%	0%	x
-C S MONO→*F,*N	2	0%	0%	0%	x

2.8.2 Historical gender changes

The second part of the sixth criterion involves determining whether there is evidence of nouns changing gender in the history of German in the direction that a constraint would predict. In order to do this, historical data taken from Ebert et al. (1993: 175f., 180, 187), who list 727 nouns whose gender has changed in some way since the Early New High German period (ca. 1350 – ca. 1650), has been analysed. The nouns listed have either changed from belonging to one gender to belonging to a different gender, or they have changed from being multiple gender nouns to belonging to just one gender. A summary of the overall changes can be seen in the table below.

⁹ See previous footnote.

	Masculine	Feminine	Neuter	Loss
Former masc.	-	209	107	-316
Former fem.	141	-	60	-201
Former neut.	133	77	-	-210
<i>Gain</i>	+274	+286	+167	-
TOTAL +/-	-42	+85	-43	

Table adapted from Watts (2016: 3)

In total, evidence for 66 constraints was found, in that there are at least two nouns in the list which have changed gender in the direction that the constraint would predict. These constraints are shown in the table below.

Constraint	Net gain/loss of nouns			Predicted direction?	Examples
	M	F	N		
-E→*M,*N	-132	+168	-36	✓	<i>Blume (M>F), Ratte (M>F)</i>
SUFFIX.-E→*M,*N	-47	+66	-19	✓	<i>Taufe (M>F), Miete (N>F)</i>
MONOSYLL.→*F,*N	+59	-19	-40	✓	<i>Brief (F>M), Dienst (N>M)</i>
DEVERBAL STEM→*F,*N	+35	-16	-19	✓	<i>Tausch (F>M), Trank (N>M)</i>
GE→*M,*F	-16	-8	+24	✓	<i>Gemach (M>N), Geschöpf (F>N)</i>
-N(C)(C) MONO→*F,*N	+17	-2	-15	✓	<i>Dunst (F>M), Fund (N>M)</i>
-NIS→*M,*F; -NIS→*M	0	-15	+15	✓	<i>Ereignis (F>N), Bündnis (F>N)</i>
CVC MONO→*F	+13	-8	-5	✓	<i>Bach, Ton, Schoß (F>M)</i>
CVCC MONO→*N	+13	-2	-11	✓	<i>Duft, Nerv, Schacht (F>M)</i>
CVCC MONO→*F	+13	-2	-11	✓	<i>Docht, Fund, Wurm (N>M)</i>
CCVC MONO→*F,*N	+10	-2	-8	✓	<i>Brief, Fluch, Flug (F>M)</i>
-C S MONO→*F,*N	+9	-6	-3	✓	<i>Flachs (N>M), Gips (N>M)</i>
D- MONO→*F	+9	-6	-3	✓	<i>Duft (F>M), Durst (F>M)</i>
-ER→*F,*N	+9	-7	-2	✓	<i>Zucker (N>M), Flitter (F>M)</i>
-EN→*F,*N; -EN→*M,*F; -EN→*F	+7	-16	+9	✓	<i>Schatten (F>M), Leben (M>N)</i>
VV _{DIPHTH} C MONO→*F	+8	-2	-6	✓	<i>Laut (F>M), Keim (F>M)</i>
NO INFO.→*M,*F	-6	-1	+7	✓	<i>Ding (M>N), Zeug (M>N)</i>
[]C- MONO→*F,*N	+7	-3	-4	✓	<i>Schmelz (N>M)</i>
-FT/-CHT→*M,*N	-2	+7	-5	✓	<i>Schrift (N>F), Luft (M>F)</i>
CVCCC MONO→*F,*N	+6	-5	-1	✓	<i>Dunst, Pelz, Durst (F>M)</i>
SUP.→*M,*F	-7	+1	+6	✓	<i>Besteck (M>N), Getreide (M>N)</i>
T- MONO→*F	+6	-2	-4	✓	<i>Ton (F>M), Talg (N>M)</i>
V _{SHORT STOP C} →*F,*N	+5	-2	-3	✓	<i>Platz (F>M), Klotz (N>M)</i>
-[]C MONO→*F,*N	+5	-3	-2	✓	<i>Talg (N>M), Schild (N>M)</i>
VER-→*F	+5	-5	0	✓	<i>Verstand (N>M), Verein (N>M)</i>
VER- NO SUFF.→*F	+5	-1	-4	✓	<i>Verkehr (N>M), Vertrag (N>M)</i>
-TUM→*M,*F; -TUM→*F	-3	-2	+5	✓	<i>Wachstum (M>N)</i>
LIVING/WORKING→*M,*F	-1	-3	+4	✓	<i>Schiff (F>N)</i>
-ENZ→*M,*N	-3	+4	-1	✓	<i>Pestilenz (N>F), Reverenz (M>F)</i>
FEAR→*M,*N	-3	+4	-1	✓	<i>Angst (M>F), Not (M>F)</i>
-ÜR/-ÜHR /y:r/→*M,*N	-2	+4	-2	✓	<i>Ungebühr, Willkür (M>F)</i>
TREES→*M,*N	-4	+4	0	✓	<i>Zeder (M>F), Esche (N>F)</i>
-EL→*F,*N	+4	-2	-2	✓	<i>Scheitel (F>M)</i>
CV MONO→*N	+4	0	-4	✓	<i>Tee, Nu (N>M)</i>
-IG/-ICH→*F,*N	+4	-2	-2	✓	<i>Bottich (F>M)</i>
-ST/-SCHT/-ZT→*N; -ST/-SCHT/-ZT→*F,*N	+4	-1	-3	✓	<i>Ernst (N>M)</i>
-UM→*M,*F	-3	0	+3	✓	<i>Datum (M>N)</i>

-ETT→*M,*F	-2	-1	+3	✓	<i>Ballett (M>N)</i>
-L MONO→*F	+3	-1	+2	✓	<i>Stall (N>M), Öl (M>N)</i>
-(C)(C)/[MONO→*F,*N	+3	-3	0	✓	<i>Dorsch (F>M)</i>
CCVC MONO→*F,*N	+3	-2	-1	✓	<i>Strahl (F>M), Zweig (N>M)</i>
-AL→*M,*F	-2	0	+2	✓	<i>Spital (M>N)</i>
-LT MONO→*M,*F	-1	-1	+2	✓	<i>Zelt (F>N), Pult (M>N)</i>
-EKT→*M,*F; -EKT→*F	-2	0	+2	✓	<i>Projekt (M>N)</i>
DOM. ANIMALS→*M,*F	-1	-1	+2	✓	<i>Pferd (M>N), Schwein (F>N)</i>
-SAL→*M; -SAL→*M,*F	-2	0	+2	✓	<i>Irrsal, Scheusal (M>N)</i>
CHEM COMP/SUBS→*M,*F	-2	0	+2	✓	<i>Zink (M>N), Messing (M>N)</i>
METALS→*M,*F	-2	0	+2	✓	<i>Zink (M>N), Messing (M>N)</i>
INSTITUTIONS→*M,*F	-1	-1	+2	✓	<i>Gefängnis (F>N), Spital (M>N)</i>
-AGE→*M,*N	0	+2	-2	✓	<i>Courage (N>F)</i>
MUS. INSTR.→*M,*N; MUS. INSTR.→*M	-1	+2	-1	✓	<i>Violine (N>F)</i>
FLOWERS→*M,*N	-2	+2	0	✓	<i>Lilie (M>F)</i>
CONDUCT→*M,*N	-1	+2	-1	✓	<i>Sitte (M>F), Moral (N>F)</i>
GROUPS→*M,*N	-2	+2	0	✓	<i>Kommune (N>F)</i>
R- MONO→*F	+2	0	-2	✓	<i>Rauch (N>M)</i>
CCVV _{DIPHTHONG} - MONO→*F	+2	-1	-1	✓	<i>Pfau (F>M), Schrei (N>M)</i>
FISH→*F,*N	+2	-2	0	✓	<i>Dorsch (F>M), Karpfen (F>M)</i>
CCVCC→*F,*N	+2	0	-2	✓	<i>Grund (N>M)</i>

The constraints in the above table can therefore be classed as diachronically productive in accordance with Enger's (2009) definition of the term. These constraints thus satisfy the second part of the sixth criterion.

The evidence gathered also suggests that the following constraints are not diachronically productive. This is because a number of nouns can be seen to have changed gender in a different direction to the one which the principle dictates. These constraints therefore do not satisfy the second part of the sixth criterion.

Constraint	Net gain/loss of nouns			Predicted direction?	Examples
	M	F	N		
-NIS→*M,*N	0	-15	+15	x	<i>Ereignis (F>N)</i>
-ER→*M,*F; -ER→*M,*N	+9	-7	-2	x	<i>Zucker (N>M)</i>
SHEET/STRIP/SLAB→*M,*F	-7	+6	+1	x	<i>Kachel (M>F)</i>
-TUM→*F,*N	-3	-2	+5	x	<i>Wachstum (M>N)</i>
-EL→*M,*F; -EL→*M,*N	+4	-2	-2	x	<i>Scheitel (F>M)</i>
PLANTS→*F,*N	-2	+3	-1	x	<i>Kresse (M>F)</i>
NAUTICAL→*M,*N	0	-2	+2	x	<i>Heck (F>N), Klafter (F>N)</i>
-EKT→*F,*N	-2	0	+2	x	<i>Projekt (M>N)</i>
-SAL→*M,*N	-2	0	+2	x	<i>Irrsal, Scheusal (M>N)</i>
V _{LONG} C MONO→*F	0	+2	-2	x	<i>Naht (N>F)</i>
MUS. INSTR.→*M,*F	-1	+2	-1	x	<i>Violine (N>F)</i>

In sum, it has been established that the sixth criterion is of great value, since evidence of productivity indicates the presence of a constraint in a native speaker's gender assignment system. In terms of the first part of the criterion, 64 of the constraints were found to be productive and 15 were found to be unproductive. In terms of the second part, 66 of the constraints were found to be productive and 11

were found to be unproductive. There is some overlap between the results for the first and second parts, but many of the constraints were only able to be deemed productive/unproductive according to one part.

For the remaining constraints, there was simply not enough evidence by which to assess their productivity. That is to say, there were either no nouns at all or just one noun with the feature invoked by the constraint, which does not constitute sufficient evidence by which to judge productivity. This is largely due to the relatively small samples used to evaluate the constraints, which also means that many of the judgements which were made about the constraints were based on relatively small numbers of nouns. Consequently, the conclusions drawn should be seen as an indication of productivity only.

2.9 Refinement of constraint set

It is now necessary to use the findings from 2.3 – 2.8 to determine which GENDER FEATURES constraints are credible and will therefore be used in the analysis in Chapter 3, and which of the constraints lack independent evidence and should therefore be eliminated from the constraint set. Enger (2009: 1294) does not specify exactly which criteria he believes a constraint must fulfil in order to be deemed valid: ‘there is no hard and fast answer as to which [constraints] should be accepted and which should not’. He does, however, give some indication of what might constitute sufficient evidence for a constraint.

Firstly, Enger (2009: 1294) states that a constraint which fares well on all criteria is more plausible than a constraint which fulfils just one criterion. He claims that we should be sceptical of constraints which have a low score, but argues that it is important not to dismiss a constraint without good reason. However, if a constraint does not meet any of the criteria then we can ‘disregard it as a serious contender’ (Enger 2009: 1294).

In line with this, the following constraints will be eliminated from the constraint set, since they failed all six criteria.

ANNOUNCE→*M,*F	INTEGRATED PARTS→*F,*N	SEMANTIC DIMINUTIVES→*M,*N
-EKT→*M,*F	IMPROPER SUPER-	SEPARABLE PARTS→*M,*F
ELAB. INSTR.→*M,*N	ORDINATES→*M,*F	SUPERORDINATES OF IN-
EXTROVERSION→*F,*N	OPEN WATER→*M,*N	DIRECT REFERENCE→*M,*F
-IV UNSTRESSED→*F,*N	PCC GARMENT→*M,*F	TIME: SHORT→*F,*N

Conversely, the following constraints will be included in the analysis in Chapter 3, since they met all six criteria.

-AT→*M,*F	-EI→*M,*N	-IE [‘i:’]→*M,*N	-UNG→*M,*N
-CHEN→*M,*F	-ER SUFF.→*F,*N	-IN SUFF.→*M,*N	-UM→*M,*F
CONVERSION→*M,*F	GE→*M,*F	-LING→*F,*N	WOMEN→*M,*N
DEVERBAL STEM→*F,*N	GEN. PERSON→*F,*N	MEN→*F,*N	
-E→*M,*N	-HEIT/-(IG)KEIT→*M,*N	SUFFIX.-E→*M,*N	

For the remaining constraints, the results are not so clear cut, and so the relative importance of each of the criteria will need to be examined more closely. To start with the first criterion (coverage of candidate nouns, Section 2.3), satisfaction of this criterion is key, given that if a principle is unable to account for a “reasonable share” of candidate nouns, then it is both of little practical use and unlikely to be picked up on by native speakers. Moreover, this justification is the basis upon which almost all gender assignment principles in the literature have been proposed (e.g. Zubin & Köpcke 1981, 1984, 1986; Köpcke 1982; Mills 1986; Wegener 1995; Nelson 1998 etc.), and one whose importance has been explicitly recognised by e.g. Mills (1986: 114), Heringer (1995: 212) and Wegener (2000: 514).

For the second criterion (typological parallels, Section 2.4), on the other hand, we have seen that, while typological support for a constraint can be extremely valuable (e.g. Kager 1999: 1, Hayes 2004: 291, McCarthy 2008: 212), Enger (2009: 1291) asserts that language-specific generalisations cannot be excluded a priori, and therefore fulfilment of the second criterion is desirable but not essential. This is especially important given the methodological limitations of 2.4, namely the investigation of three languages only and the lack of verification of the principles proposed. Similarly, for the third criterion (parallels within the language, Section 2.5), despite its value (highlighted by e.g. Mills (1986: 114) and Corbett (1991: 31)), Enger (2009: 1292) also states that its fulfilment cannot be a requirement, since ‘the possibility of rules based on “unique” features cannot be ruled out *a priori*’.

Fulfilment of Enger’s (2009: 1292) fourth criterion (recognisability and semantic clarity, Section 2.6), however, is crucial for constraints involving semantic features, because if a native speaker is unable to recognise a feature, then the feature cannot practically be used in gender assignment. It is also unlikely that a native speaker would be able to acquire a constraint which involved an unclear and unrecognisable feature. The necessity for GENDER FEATURES constraints to be clear is also supported by e.g. Comrie (1999: 461).

As established in 2.7, fulfilment of Enger’s (2009: 1293) fifth criterion (accounting for exceptions) is useful, given that if a constraint is able to account for exceptions to another constraint then this indirectly indicates that the constraint is valid. However, as stated in 2.7, this criterion is best seen as additional evidence for a constraint, rather than as proof of its existence, given that accounting for exceptions alone is not sufficient evidence for a constraint, and, moreover, is the basis upon which many dubious constraints are likely to have been proposed.

Finally, satisfaction of Enger's (2009: 1294) sixth criterion (productivity, Section 2.8) is important, given that it is evidence of the psychological reality of a constraint, and therefore an indication of the constraint's validity. Enger (2009: 1294) even claims that it would be 'reasonable' to make fulfilment of this criterion a requirement for any constraint. However, due to size of the corpora upon which the productivity of the constraints was tested, satisfaction of this criterion cannot be made essential. For the majority of constraints tested, there was simply not sufficient evidence to make a judgement about their productivity. For example, otherwise plausible constraints such as LANGUAGES→*M,*F and ROCKS/MINERALS→*F,*N meet all criteria apart from the sixth, since insufficient evidence was found. It therefore cannot be expected that all credible constraints show diachronic productivity in 2.8.

In summary, it has been established that fulfilment of the first and fourth (where applicable) criteria is necessary for any constraint that will appear in the analysis in Chapter 3, and that criteria two, three, five and six all provide useful additional evidence for a constraint, but, for various reasons, cannot individually be classed as essential. We can therefore reasonably demand that a constraint meets criteria one and four, and at least one other criterion out of the rest. Satisfaction of a minimum of three criteria is also supported by Enger (2009: 1294).

The following constraints will therefore be eliminated from the constraint set, since they failed to meet criteria one and four and at least one other.

-AL→*M,*F	-IT→*M,*F	BIRDS→*F,*N	MUS. COMP.→*M,*F
-AR SUFF.→*F,*N	-IV→*F,*N	BOATS→*F,*N	MUS. INSTR.→*M,*F
-AR SUFF.→*M,*F	KR- MONO→*F,*N	BODIES OF WATER→*N	MUS. INSTR.→*M,*N
-AT→*F,*N	-[I]C MONO→*F,*N	CONDUCT→*M,*N	MUS. TERMS→*M,*F
-EE/-ÉE/-É→*F,*N	-LT MONO→*M,*F	DISEASES→*M,*N	NAUTICAL→*M,*N
-EE/-ÉE/-É→*M,*F	MONOSYLL.→*F,*N	DOM. ANIMALS→*M,*F	NO INFO.→*M,*F
-EE/-ÉE/-É→*M,*N	-NIS→*M,*F	DWELL→*M,*F	OPENINGS→*M,*F
-EHR/-ÄHR→*M,*N	-NIS→*M,*N	F. HOLLOW→*M,*F	ORGANS→*M,*F
-EKT→*F,*N	-O→*M,*F	FEAR→*M,*N	PICTURES→*M,*F
-EL→*F,*N	-ON→*M,*F	GEN. PERSON→*M,*F	PLANTS→*F,*N
-EL→*M,*F	-RIS→*M,*N	GESTURES→*M,*N	POWER/STRENGTH→*M,*N
-EL→*M,*N	-SAL→*M,*N	GROUPS→*M,*F	PREP. DISHES→*M,*F
-EN→*F,*N	-ST/-SCHT/-ZT→*F,*N	HUNTING→*M,*N	RIVERS EUROPE→*M,*N
-EN→*M,*F	TR→*F,*N	INSTITUTIONS→*M,*F	SHARP→*M,*N
-ENT→*M,*F	-/tʃ/;/Ntʃ/;/Nj/ MONO→*F,*N	INTROVERSION→*M,*N	SHEET/STRIP/SLAB→*M,*F
-FT/-CHT→*M,*N	-TUM→*F,*N	LIVING/WORKING→*M,*F	STICK→*F,*N
-IER /i:r/ POLY.→*M,*F	VCC MONO→*F,*N	MAMMALS→*F,*N	TEMPERATURES→*M,*N
-IG→*F,*N	ANG. SHAPES→*M,*F	MESH→*M,*F	UNITS OF TIME→*F,*N

The constraints below, on the other hand, all passed the first, fourth and at least one other criterion, and will thus be included in the constraint set for the analysis in Chapter 3.

-A→*M,*N	-ER→*F,*N	-N(C)(C) MONO→*F,*N	VER- NO SUFF.→*F	MON. UNITS→*F,*N
-ADE→*M,*N	-ETT→*M,*F	-NIS→*M	V _{LONG} MONO→*F	MOUNTAINS→*F,*N
-AGE→*M,*N	-ETTE→*M,*N	-(O)LOGE→*F,*N	V _{LONG} C MONO→*F	MUS. INSTR.→*M
-ANER→*F,*N	-EUR/ÖR→*F,*N	-ON UNSTRESS.→*M,*F	V _{SHORT} STOP C→*F,*N	MUS. NOTES→*M,*F
-ANG→*F,*N	-EUR/ÖR→*N	-OR→*F,*N	V _{STRESSED} [I]→*M,*F	NAME: MAN→*F,*N

-ANT→*F,*N	-IV STRESSED→*M,*F	R- MONO→*F	V _{STRESSED} [m]→*M,*F	NAME: WOMAN→*M,*N
-ANZ→*M,*N	-IE [iə]→*M,*N	-/r/ STOP (C) MONO→*F,*N	VV _{DIPHTH,C} MONO→*F	REGIONS→*M,*F
-B→*F	-IG/-ICH→*F,*N	-SAL→*M,*F	ALC. DRINKS→*F,*N	RIVERS NON-GER.→*F,*N
-C S MONO→*F,*N	-IK→*M,*N	-SAL→*M	CHEM COMP./	ROCKS/MINERALS→*F,*N
CCCVC MONO→*F,*N	-IN [i:n]→*M,*F	-[ɨ] MONO→*F,*N	SUBS.→*M,*F	LEATHER TYPES→*M,*F
CCVC MONO→*F,*N	-INE→*M,*N	[[ɨ]C- MONO→*F,*N	CITIES/TOWNS→*M,*F	RIVERS NON-
CCVCC→*F,*N	-ING→*M,*F	-SCHAFT→*M,*N	COUNTRIES→*M,*F	EUROPE→*F,*N
CCVV _{DIPHTHONG} MONO→*F	-ION→*M,*N	-SKOP→*M,*F	DAYS→*F,*N	SCI. UNITS→*M,*F
CV MONO→*N	-IS→*M,*N	-ST/-SCHT/-ZT→*N	ELEMENTS→*M,*F	SEASONS→*F,*N
CVC MONO→*F	-ISMUS→*F,*N	T- MONO→*F	FABRIC→*F,*N	SUP.→*M,*F
CVCC MONO→*F	-IST→*F,*N	-TÄT→*M,*N	FISH→*F,*N	TIMES OF DAY→*F,*N
CVCC MONO→*N	-IT→*F	-THEK→*M,*N	FLOWERS→*M,*N	TREES→*M,*N
CVCCC MONO→*F,*N	-IT→*F,*N	-TUM→*F	FRUITS→*M,*N	VEG.ROOT→*M,*N
D- MONO→*F	-ITIS→*M,*N	-TUM→*M,*F	GAMES→*M,*F	WASTE→*F,*N
DEVERBAL -T→*M,*N	-IUM→*M,*F	-UR SUFF.→*M,*N	GROUPS→*M,*N	WORD CLASSES→*M,*F
-EDER→*M,*F	-IV→*F	-UR STRESSED→*M,*N	ISLANDS→*M,*F	YOUNG ANIMALS→*M,*F
-EKT→*F	-IV→*M,*F	-ÜR/-ÜHR /y:r/→*M,*N	LANGUAGES→*M,*F	
-EN→*F	-L MONO→*F	-US→*F,*N	LETTERS→*M,*F	
-ENT→*F,*N	-MA→*M,*F	VC MONO→*M,*F	MALE ANIMALS→*F,*N	
-ENZ→*M,*N	-MENT→*M,*F	VCCC MONO→*M	METALS→*M,*F	

For a full summary of the results in Chapter 2, see Appendix E.

Constraints with multiple versions

The second question which ought to be considered in this section is what is to be done with the constraints of which there are multiple versions, e.g. GEN. PERSON→*F,*N and GEN. PERSON→*M,*F (introduced in 2.2). In some cases, it is clear which version is best and should therefore be included in the constraint set, since one fares significantly better by Enger's (2009) criteria than the other(s). Examples of this can be seen below.

Constraint	1	2	3	4	5	6	TOTAL
GEN. PERSON→*M,*F	x	✓	✓	✓	x	x	3
☞ GEN. PERSON→*F,*N	✓	✓	✓	✓	✓	✓	6
MUS. INSTR.→*M,*F	x	✓	x	✓	x	x	2
MUS. INSTR.→*M,*N	x	✓	x	✓	x	-	2
☞ MUS. INSTR.→*M	✓	✓	x	✓	x	✓	4
GROUPS→*M,*F	x	x	✓	✓	x	x	2
☞ GROUPS→*M,*N	✓	x	✓	✓	✓	✓	5
-AT→*F,*N	x	✓	✓	n/a	x	x	2
☞ -AT→*M,*F	✓	✓	✓	n/a	✓	✓	5
-EKT→*M,*F	x	x	x	n/a	x	-	0
-EKT→*F,*N	✓	x	x	n/a	x	x	1
☞ -EKT→*F	✓	x	x	n/a	x	✓	2
-EN→*F,*N	x	x	✓	n/a	✓	✓	3
-EN→*M,*F	x	x	✓	n/a	✓	✓	3
☞ -EN→*F	✓	x	✓	n/a	✓	✓	4

-ER→*M,*F	x	x	✓	n/a	x	x	1
-ER→*M,*N	x	x	✓	n/a	x	x	1
☞ -ER→*F,*N	✓	x	✓	n/a	✓	✓	4

-ENT→*M,*F	x	✓	✓	n/a	x	-	2
☞ -ENT→*F,*N	✓	✓	✓	n/a	x	-	3

-ST/-SCHT/-ZT→*F,*N	x	x	✓	n/a	x	✓	2
☞ -ST/-SCHT/-ZT→*N	✓	x	✓	n/a	x	✓	3

-NIS→*M,*N	x	✓	✓	n/a	x	x	2
-NIS→*M,*F	x	✓	✓	n/a	x	✓	3
☞ -NIS→*M	✓	✓	✓	n/a	x	✓	4

In other cases, all versions of the constraint fail according to the criteria, thereby excluding the constraint from the analysis entirely.

Constraint	1	2	3	4	5	6	TOTAL
-EL→*M,*F	x	x	✓	n/a	✓	x	2
-EL→*F,*N	x	x	✓	n/a	✓	-	2
-EL→*M,*N	x	x	✓	n/a	✓	-	2

-EE/-ÉE/-É→*M,*F	x	x	✓	n/a	x	-	1
-EE/-ÉE/-É→*F,*N	x	x	✓	n/a	x	-	1
-EE/-ÉE/-É→*M,*N	x	x	✓	n/a	x	-	1

There are also, however, six instances in which two versions of a constraint appear equally (or almost equally) plausible according to Enger's (2009) criteria. In such cases, the only way to determine which version of the constraint is the most plausible is to test each one on the noun sample. This will be done in Section 3.2.1.

Constraint	1	2	3	4	5	6	TOTAL
☞ CVCC MONO→*F	✓	x	x	n/a	✓	✓	3
☞ CVCC MONO→*N	✓	x	x	n/a	✓	✓	3

☞ -EUR/ÖR→*F,*N	✓	x	✓	n/a	x	-	2
☞ -EUR/ÖR→*N	✓	x	✓	n/a	x	-	2

-TUM→*F,*N	x	✓	✓	n/a	x	x	2
☞ -TUM→*M,*F	✓	✓	✓	n/a	x	✓	4
☞ -TUM→*F	✓	✓	✓	n/a	x	✓	4

☞ -SAL→*M	✓	x	✓	n/a	x	✓	3
☞ -SAL→*M,*F	✓	x	✓	n/a	x	✓	3
-SAL→*M,*N	x	x	✓	n/a	x	-	1

-IV→*F,*N	x	✓	✓	n/a	x	-	2
☞ -IV→*M,*F	✓	✓	✓	n/a	✓	-	4
☞ -IV→*F	✓	✓	✓	n/a	✓	-	4

-IT→*M,*F	x	x	✓	n/a	✓	-	2
☞-IT→*F,*N	✓	x	✓	n/a	x	-	2
☞-IT→*F	✓	x	✓	n/a	✓	-	3

Revising failed constraints

The third and final question to be considered in this section is, for constraints which failed the first criterion (share of candidate nouns, Section 2.3), and have therefore been excluded from the constraint set, yet fare well by the remaining criteria such that their exclusion rests solely upon their failure of the first criterion, whether an amended version of the constraint whereby it eliminates just one gender (as opposed to two) would emerge as valid according to the criteria (criterion one in particular).

For instance, the constraint -AL→*M,*F has typological parallels and shows evidence of productivity, yet only covers 53% of candidate nouns, thereby failing the first criterion. However, if the constraint were modified into -AL→*F, eliminating just the possibility of being feminine, then it would cover 99% of candidate nouns, as well as still having typological parallels and showing evidence of productivity. It therefore seems worth considering modified versions of such constraints rather than omitting them entirely.

It is important to note that several principles of this type have already been proposed in the literature, such as those listed below (see e.g. Altmann & Raettig 1973, Köpcke 1982, Steinmetz & Rice 1989, Köpcke & Zubin 1996, Nelson 1998, Steinmetz 2006 etc.).

D- MONO→ F	-IV→*F	R- MONO→*F	-TUM→*F
-EN→*F	-L MONO→*F	-ST/SCHT/ZT→*N	VER- NO SUFF.→*F
-IT→*F	-NIS→*M	T- MONO→*F	VV _{DIPHTH} C MONO→*F

The table below shows the constraints which failed the first criterion but otherwise would have been deemed plausible according to the criteria (column 1), along with the possible variants of these constraints and the percentage coverage of each.

Original Constraint	%	% required by TP (incl. 10% margin)	Variant 1 *M	%	Variant 2 *F	%	Variant 3 *N	%
-EL→*F,*N	45	74	-EL→*M	55	-EL→*F	72	-EL→*N	72
MUS. TERMS→*M,*F	48	70	MUS. TERMS→*M	85	MUS. TERMS→*F	63	MUS. TERMS→*N	52
INSTITUTIONS→*M,*F	38	64	INSTITUTIONS→*M	89	INSTITUTIONS→*F	49	INSTITUTIONS→*N	62
-EE/-ÉE/-É→*F,*N	33	70	-EE/-ÉE/-É→*F	75	-EE/-ÉE/-É→*F	67	-EE/-ÉE/-É→*N	67
MUS. COMP.→*M,*F	44	70	MUS. COMP.→*M	87	MUS. COMP.→*F	57	MUS. COMP.→*N	66
DWELL→*M,*F	48	66	DWELL→*M	84	DWELL→*F	64	DWELL→*N	52
-AL→*M,*F	53	70	-AL→*M	55	-AL→*F	99	-AL→*N	57

-O→*M,*F	49	73	-O→*M	55	-O→*F	94	-O→*N	51
-IER /i:r/ POLY.→*M,*F	41	64	-IER /i:r/ POLY.→*M	49	-IER /i:r/ POLY.→*F	92	-IER /i:r/ POLY.→*N	59
PLANTS→*F,*N	39	68	PLANTS→*M	61	PLANTS→*F	52	PLANTS→*N	89
RIVERS EUROPE→*M,*N	56	67	RIVERS EUROPE→*M	44	RIVERS EUROPE→*F	44	RIVERS EUROPE→*N	100
PICTURES→*M,*F	58	61	PICTURES→*M	88	PICTURES→*F	70	PICTURES→*N	42
DOM. ANIMALS→*M,*F	52	59	DOM. ANIMALS→*M	80	DOM. ANIMALS→*F	72	DOM. ANIMALS→*N	48
ORGANS→*M,*F	29	55	ORGANS→*M	82	ORGANS→*F	53	ORGANS→*N	71
DISEASES→*M,*N	65	68	DISEASES→*M	78	DISEASES→*F	35	DISEASES→*N	87
-FT/-CHT→*M,*N	66	67	-FT/-CHT→*M	75	-FT/-CHT→*F	34	-FT/-CHT→*N	90
-LT MONO→*M,*F	28	61	-LT MONO→*M	40	-LT MONO→*F	88	-LT MONO→*N	72
MONOSYLL.→*F,*N	64	76	MONOSYLL.→*M	36	MONOSYLL.→*F	86	MONOSYLL.→*N	78
MAMMALS→*F,*N	60	66	MAMMALS→*M	40	MAMMALS→*F	87	MAMMALS→*N	74
BIRDS→*F,*N	64	66	BIRDS→*M	36	BIRDS→*F	69	BIRDS→*N	94
STICK→*F,*N	59	65	STICK→*M	41	STICK→*F	73	STICK→*N	86
UNITS OF TIME→*F,*N	49	64	UNITS OF TIME→*M	51	UNITS OF TIME→*F	71	UNITS OF TIME→*N	78
-[I]C MONO→*F,*N	67	68	-[I]C MONO→*M	33	-[I]C MONO→*F	88	-[I]C MONO→*N	80

As can be seen in the above table, all but one of the constraints (-EL→*F,*N) have at least one variant which is considered valid according to Enger's (2009) criteria. Around half have two plausible variants. For those with just one plausible version, this version will be included in the constraint set for the analysis in Chapter 3, since it passes criteria one, four and at least one other. For those with two plausible versions, since these are deemed equally plausible according to Enger's (2009) criteria, the most plausible version will have to be identified through the testing of each one on the noun sample, as with the similar cases above. This will be done in 3.2.1.

2.10 Summary of results

The table below provides a summary of the key results from this chapter. All of the constraints tested have been categorised according to their plausibility as determined by Enger's (2009) six criteria. Those which are deemed plausible (i.e. pass criteria one, four and at least one other) will form part of the constraint set for the analysis in Chapter 3, and those which are deemed implausible (i.e. do not pass criteria one, four and at least one other) will be removed from the constraint set.

Plausible		Implausible
-A→*M,*N	VC MONO→*M,*F	ANG. SHAPES→*M,*F
-ADE→*M,*N	VCCC MONO→*M	ANNOUNCE→*M,*F
-AGE→*M,*N	VER- NO SUFF.→*F	BIRDS→*F,*N
-AL→*F	V _{STRESSED} [I]→*M,*F	BOATS→*F,*N
-AN→*M,*F	V _{STRESSED} [m]→*M,*F	BODIES OF WATER→*N
-ANER→*F,*N	V _{LONG} MONO→*F	CONDUCT→*M,*N
-ANG→*M,*F	V _{LONG} C MONO→*F	DISEASES→*M,*N
-ANT→*F,*N	V _{SHORT} STOP C→*F,*N	DOM. ANIMALS→*M,*F
-ANZ→*M,*N	VV _{DIPHTH} C MONO→*F	DWELL→*M,*F
-AT→*M,*F	ACAD. DISCIPLINES→*M,*N	ELAB. INSTR.→*M,*N
-AR→*M,*F	ALC. DRINKS→*F,*N	EXTROVERSION→*F,*N

-B→*F	CHEESES→*F,*N	F. HOLLOW→*M,*F
-C S MONO→*F,*N	CHEM COMP./SUBS.→*M,*F	FEAR→*M,*N
-(C)(C)/J/ MONO→*F,*N	CITIES/TOWNS→*M,*F	GEN. PERSON→*M,*F
CCVC MONO→*F,*N	COUNTRIES→*M,*F	GESTURES→*M,*N
CCVC MONO→*F,*N	REGIONS→*M,*F	GROUPS→*M,*F
CCVCC→*F,*N	DAYS→*F,*N	HUNTING→*M,*N
CCVV _{DIPHTHONG} - MONO→*F	DWELL→*M	IMPROPER SUPERORDINATES→*M,*F
-CHEN→*M,*F	ELEMENTS→*M,*F	INSTITUTIONS→*M,*F
CONVERSION→*M,*F	FABRIC→*F,*N	INTEGRATED PARTS→*F,*N
CV MONO→*N	FISH→*F,*N	INTROVERSION→*M,*N
CVC MONO→*F	FLOWERS→*M,*N	LIVING/WORKING→*M,*F
CVCCC MONO→*F,*N	FRUITS→*M,*N	MAMMALS→*F,*N
D- MONO→*F	GAMES→*M,*F	MESH→*M,*F
DR-→*F,*N	GEN. PERSON→*F,*N	MUS. INSTR.→*M,*F
DEVERBAL STEM→*F,*N	GROUPS→*M,*N	MUS. INSTR.→*M,*N
DEVERBAL -T→*M,*N	INSTITUTIONS→*M	MUS. TERMS→*M,*F
-E→*M,*N	INSECTS→*M,*N	NAUTICAL→*M,*N
-EE/-ÉE/-É→*F	ISLANDS→*M,*F	NO INFO.→*M,*F
-EDER→*M,*F	LANGUAGES→*M,*F	OPEN WATER→*M,*N
-EKT→*F	LEATHER TYPES→*M,*F	OPENINGS→*M,*F
-EN→*F	LETTERS→*M,*F	ORGANS→*M,*F
-ENT→*F,*N	MALE ANIMALS→*F,*N	PCC GARMENT→*M,*F
-ENZ→*M,*N	MEN→*F,*N	PICTURES→*M,*F
-ER→*F,*N	METALS→*M,*F	PLANTS→*F,*N
-(ER)EI→*M,*N	MON. UNITS→*F,*N	POWER/STRENGTH→*M,*N
-ER SUFF.→*F,*N	MOUNTAINS→*F,*N	PREP. DISHES→*M,*F
-ETT→*M,*F	MUS. COMP.→*M	SEMANTIC DIMINUTIVES→*M,*N
-ETTE→*M,*N	MUS. TERMS→*M	SEPARABLE PARTS→*M,*F
GE→*M,*F	MUS. INSTR.→*M	SHARP→*M,*N
-HEIT/- (IG)KEIT→*M,*N	MUS. NOTES→*M,*F	SHEET/STRIP/SLAB→*M,*F
-IE [i:]→*M,*N	MUSIC TYPES→*F,*N	SPEECH→*M,*N
-IE [iə]→*M,*N	NAME: MAN→*F,*N	STICK→*F,*N
-IER /i:r/ POLY.→*F	NAME: WOMAN→*M,*N	SUPERORDINATES OF INDIRECT REFERENCE→*M,*F
-IG/-ICH→*F,*N	PLANTS→*N	TEMPERATURES→*M,*N
-IK→*M,*N	PRIMATES→*F,*N	TIME: SHORT→*F,*N
-ILLE→*M,*N	RIVERS EUROPE→*N	UNITS OF TIME→*F,*N
-INER→*F,*N	RIVERS NON-GER.→*F,*N	-AL→*M,*F
-IN [i:n]→*M,*F	RIVERS NON-EUROPE→*F,*N	-AR SUFF.→*F,*N
-IN SUFF.→*M,*N	ROCKS/MINERALS→*F,*N	-AR SUFF.→*M,*F
-INE→*M,*N	SCI. UNITS→*M,*F	-AT→*F,*N
-ING→*M,*F	SEASONS→*F,*N	-EE/-ÉE/-É→*F,*N
-ION→*M,*N	SUP.→*M,*F	-EE/-ÉE/-É→*M,*F
-IS→*M,*N	TIMES OF DAY→*F,*N	-EE/-ÉE/-É→*M,*N
-ISMUS→*F,*N	TREES→*M,*N	-EHR/-ÄHR→*M,*N
-IST→*F,*N	VEG.ROOT→*M,*N	-EKT→*F,*N
-ITIS→*M,*N	WASTE→*F,*N	-EKT→*M,*F
-IUM→*M,*F	WOMEN→*M,*N	-EL→*F,*N
-IV STRESSED→*M,*F	WORD CLASSES→*M,*F	-EL→*M,*F
-IVE→*M,*N	YOUNG ANIMALS→*M,*F	-EL→*M,*N
-L MONO→*F	CVCC MONO→*F/CVCC MONO→*N	-EN→*F,*N
-LING→*F,*N	-EUR/ÖR→*F,*N/-EUR/ÖR→*N	-EN→*M,*F
-MA→*M,*F	-IT→*F,*N/-IT→*F	-ENT→*M,*F
-MENT→*M,*F	-IV→*F/-IV→*M,*F	-ER→*M,*F
-NIS→*M	-SAL→*M/-SAL→*M,*F	-ER→*M,*N
-N(C)(C) MONO→*F,*N	-TUM→*M,*F/-TUM→*F	-IER /i:r/ POLY.→*M,*F
-O→*F	-[!]C MONO→*F/-[!]C MONO→*N	-IG→*F,*N
-(O)LOGE→*F,*N	-LT MONO→*F/-LT MONO→*N	

-OR→*F,*N -ON UNSTRESS.→*M,*F R- MONO→*F -/r/ STOP (C) MONO→*F,*N [[]C- MONO→*F,*N -SCHAFT→*M,*N -SKOP→*M,*F -ST/-SCHT/-ZT→*N SUFFIX.-E→*M,*N T- MONO→*F -TÄT→*M,*N -THEK→*M,*N -UM→*M,*F -UNG→*M,*N -UR SUFF.→*M,*N -UR STRESSED→*M,*N -ÜR/-ÜHR /y:r/→*M,*N -US→*F,*N ∅V _{LONG} MONO→*F	-FT/-CHT→*M/-FT/-CHT→*N BIRDS→*F/BIRDS→*N DISEASES→*M/DISEASES→*N DOM. ANIMALS→*F/ DOM. ANIMALS→*M MAMMALS→*F/MAMMALS→*N MONOSYLL.→*F/MONOSYLL.→*N ORGANS→*M/ORGANS→*N PICTURES→*F/PICTURES→*M STICK→*F/STICK→*N UNITS OF TIME→*F/UNITS OF TIME→*N	-IT→*M,*F -IV→*F,*N -IV UNSTRESSED→*F,*N KR- MONO→*F,*N -LT MONO→*M,*F -[]C MONO→*F,*N MONOSYLL.→*F,*N -NIS→*M,*F -NIS→*M,*N -O→*M,*F -ON→*M,*F -RIS→*M,*N -/tʃ;/Ntʃ;/Nj/ MONO→*F,*N TR-→*F,*N -SAL→*M,*N -ST/-SCHT/-ZT→*F,*N -TUM→*F,*N VER-→*F VCC MONO→*F,*N
--	--	--

Chapter 3: Analysis – testing OGAT

3.1 Introduction and preliminary analysis

The main aim of this chapter is to determine whether OGAT, with a constraint set that has been independently verified, can account for the genders of the 592 nouns in the sample (introduced in 2.2). This will then give us an indication of the validity of OGAT as a way of modelling German gender assignment. As stated in 1.4.5.3 and 1.4.5.4, previous studies on OGAT have only tested the theory on a very small number of nouns, and/or have included seemingly questionable, unverified GENDER FEATURES constraints in their analyses.

In Chapter 2, all GENDER FEATURES constraints from the literature which are relevant or at least potentially relevant to the noun sample were tested according to six criteria proposed by Enger (2009), which seek independent evidence for gender assignment principles. Any constraint without sufficient independent justification was eliminated from the set (see 2.9). Accordingly, it is now necessary to test OGAT on the noun sample with the refined constraint set, in order to establish whether it can successfully predict the nouns' genders.

A preliminary analysis reveals that OGAT, when using the set of independently-verified constraints, is indeed able to account for the genders of a large proportion of the noun sample. The theory therefore continues to look promising, even when only using GENDER FEATURES constraints for which independent evidence has been found. Below it will be shown that OGAT is able to account for the genders of various nouns in the sample.

Firstly, there are a number of nouns in the sample to which OGAT assigns the correct gender based on a single, independently-verified semantic constraint. For example, *Mikrofarad* correctly receives neuter due to *SCI. UNITS*→**M*,**F* “scientific units and units of measurement cannot be masculine or feminine” (e.g. Köpcke 1982: 74, Flämig 1991: 453, Eisenberg 1999: 155, Menzel 2004: 63, Thomoglou 2004: 34, Hoberg 2004: 106, Chan 2005: 95), and *Farsi* is correctly assigned neuter due to *LANGUAGES*→**M*,**F* “names of languages cannot be masculine or feminine” (e.g. Köpcke 1982: 74, Flämig 1991: 453, Köpcke & Zubin 1996: 480, Menzel 2004: 63, Chan 2005: 96). The tableau for *Farsi* is shown in (1).

(1)

	GENDER FEATURES LANGUAGES→* <i>M</i> ,* <i>F</i>	DEFAULT HIERARCHY		
		* <i>N</i>	* <i>F</i>	* <i>M</i>
<i>Farsi</i>				*
<i>der Farsi</i>	*!			*
<i>die Farsi</i>	*!		*	
☞ <i>das Farsi</i>		*		

Similarly, there are various cases in which OGAT correctly assigns gender to nouns in the sample on the basis of a single, independently-verified formal constraint. These include *Konformismus*, assigned masculine due to $-ISMUS \rightarrow *F, *N$ (e.g. Altmann & Raettig 1973: 302, Duden Grammatik 2009: 164, Zifonun et al. 1997: 33, Engel 2009: 280, Hoberg 2004: 86, Weinrich 2007: 326), *Triolett*, assigned neuter on the basis of $-ETT \rightarrow *M, *F$ (e.g. Mills 1986: 33, Steinmetz & Rice 1989: 166, Köpcke & Zubin 1997, Hoberg 2004: 96, Menzel 2004: 68, Duden Grammatik 2009: 166), and *Legalität*, assigned feminine on the basis of $-(I)TÄT \rightarrow *M, *N$ (e.g. Flämig 1991: 453, Mills 1986: 30, Zifonun et al. 1997: 33, Hoberg 2004: 87, Menzel 2004: 66, Thomoglou 2004: 41, Weinrich 2007: 327, Duden Grammatik 2009: 165, Engel 2009: 281). The tableau for *Legalität* is given in (2).

(2)

<i>Legalität</i> 'legality'	GENDER FEATURES $-(I)TÄT \rightarrow *M, *N$	DEFAULT HIERARCHY		
		*N	*F	*M
<i>der Legalität</i>	*!			*
<i>die Legalität</i>			*	
<i>das Legalität</i>	*!	*		

OGAT is also able to account for the genders of various monosyllables in the sample, such as *Tran*, *Sog* and *Angst*,¹ which are each subject to multiple, independently-verified phonological constraints. *Tran* is correctly assigned masculine due to $-N(C)(C) \text{ MONO} \rightarrow *F, *N$ (e.g. Köpcke 1982: 100, Mills 1986: 33, Köpcke & Zubin 1997, Menzel 2004: 68, Hoberg 2004: 98), $CCVC \text{ MONO} \rightarrow *F, *N$ (Köpcke 1982: 85) and $V_{LONG}(C) \text{ MONO} \rightarrow *F$ (Zubin & Köpcke 1981: 440, Köpcke 1982: 87, 95), *Sog* is correctly assigned masculine due to $V_{LONG}(C) \text{ MONO} \rightarrow *F$ (ibid.) and $CVC \text{ MONO} \rightarrow *F$ (Köpcke 1982: 83) together with the default hierarchy (which eliminates the neuter candidate due to $*N \gg *M$), and *Angst* is correctly assigned feminine due to $VCCC \text{ MONO} \rightarrow *M$ (Köpcke 1982: 83) and $-ST/-SCHT/-ZT \rightarrow *N$ (Köpcke 1982: 97), as shown in (3).

(3)

<i>Angst</i> 'fear'	GENDER FEATURES $VCCC \text{ MONO} \rightarrow *M$ $-ST/-SCHT/-ZT \rightarrow *N$	DEFAULT HIERARCHY		
		*N	*F	*M
<i>der Angst</i>	*!			*
<i>die Angst</i>			*	
<i>das Angst</i>	*!	*		

Additionally, there are a number of nouns in the sample whose gender OGAT is able to correctly predict on the basis of independently-verified semantic and formal constraints which eliminate the same genders as each other. For instance, *Sommer* correctly receives masculine on the basis of both $SEASONS \rightarrow *F, *N$ (e.g. Flämig 1991: 452, Köpcke & Zubin 1996: 480, Hickey 2000: 629, Menzel 2004: 63,

¹ *Angst* as seen in the sample item *Examensangst* – see discussion of compounds in 3.3.3.

Chan 2005: 94, Weinrich 2007: 327, Engel 2009: 272) and $-ER \rightarrow *F, *N$ (e.g. Steinmetz 1986: 197, Wegener 1995, Hoberg 2004: 92, Eisenberg 2013: 135), and *Rubidium* correctly receives neuter on the basis of $METALS \rightarrow *M, *F$ (e.g. Köpcke 1982: 74, Menzel 2004: 63, Thomoglou 2004: 34, Weinrich 2007: 328, Eisenberg 2013: 138), $ELEMENTS \rightarrow *M, *F$ (e.g. Köpcke 1982: 74, Mills 1986: 27, Flämig 1991: 452, Köpcke & Zubin 1996: 480, Menzel 2004: 63, Thomoglou 2004: 34, Weinrich 2007: 328) and $-IUM \rightarrow *M, *F$ (e.g. Altmann & Raettig 1973: 302, Steinmetz 1985; 1986: 198, Flämig 1991: 453, Nelson 1998: 215, Menzel 2004: 66, Weinrich 2007: 326, Duden Grammatik 2009: 166), as shown in (4).

(4)

<i>Rubidium</i>	GENDER FEATURES			DEFAULT HIERARCHY		
	$-IUM \rightarrow *M, *F$	$METALS \rightarrow *M, *F$	$ELEMENTS \rightarrow *M, *F$	*N	*F	*M
<i>der Rubidium</i>	*	*	*!			*
<i>die Rubidium</i>	*	*	*!		*	
<i>☞ das Rubidium</i>				*		

There are also numerous cases of nouns in the sample to which OGAT is able to assign the correct gender whose form and meaning conflict with regard to gender assignment. Some of these cases involve OGAT resolving the conflict in favour of the gender associated with the noun's meaning (contra the form » semantics approach, see 1.4.2). For example, the noun *Porphyry* is correctly assigned masculine due to the constraints $ROCKS/MINERALS \rightarrow *F, *N$ (e.g. Köpcke 1982: 71f., Flämig 1991: 452, Weinrich 2007: 328, Engel 2009: 272) and $-ÜR/-ÜHR /y:r/ \rightarrow *M, *N$ (e.g. Mills 1986: 33, Hickey 2000: 631, Menzel 2004: 68, Chan 2005: 296) causing the neuter candidate to be eliminated, and then the default hierarchy resolving the conflict between the remaining candidates (masculine, associated with the meaning, and feminine, associated with the form) in favour of masculine due to $*F \gg *M$. A similar case is seen in *Slowake*, to which OGAT correctly assigns masculine due to a combination of the GENDER FEATURES constraints $-E \rightarrow *M, *N$ (e.g. Altmann & Raettig 1973: 302, Steinmetz 1986: 192, Wegener 1995, Köpcke & Zubin 1996: 476, Nelson 1998: 218, Rice 2006: 1396, Eisenberg 2013: 134) and $GEN. PERSON \rightarrow *F, *N$ (e.g. Köpcke 1982: 72, Mills 1986: 27, Hoberg 2004: 103, Chan 2005: 91ff., Onysko 2007: 159), and the default hierarchy, which again resolves the conflict in favour of the gender associated with the noun's meaning, i.e. masculine. This is illustrated in (5).

(5)

<i>Slowake 'Slovak'</i>	GENDER FEATURES		DEFAULT HIERARCHY		
	$GEN. PERSON \rightarrow *F, *N$	$-E \rightarrow *M, *N$	*N	*F	*M
<i>☞ der Slowake</i>		*			*
<i>die Slowake</i>	*			*!	
<i>das Slowake</i>	*	*!	*		

Conversely, there are many other cases of nouns in the sample with a conflicting form and meaning to which OGAT assigns the correct gender, where OGAT resolves the conflict in favour of the gender associated with the noun's form (contra the semantics » form approach, see 1.4.1). For example, *Seniorengymnastik* is assigned feminine by OGAT as a result of the constraints $-IK \rightarrow *M, *N$ (e.g. Altmann & Raettig 1973: 302, Zifonun et al. 1997: 33, Nelson 1998: 218, Hoberg 2004: 87, Thomoglou 2004: 41, Weinrich 2007: 327, Engel 2009: 281) and $GAMES/SPORTS \rightarrow *M, *F$ (e.g. Mills 1986: 27, Köpcke & Zubin 1996: 480, Chan 2005: 96), which eliminate the masculine candidate, and the default hierarchy, which favours feminine, the gender associated with the noun's form, over neuter, the gender associated with the noun's meaning. The tableau for *Seniorengymnastik* is shown in (6). Similarly, *Rotang* is assigned masculine as a result of the constraints $-ANG \rightarrow *F, *N$ (Altmann & Raettig 1973: 302) and $TREES \rightarrow *M, *N$ (e.g. Flämig 1991: 452, Hickey 2000: 629, Menzel 2004: 63, Thomoglou 2004: 34, Chan 2005: 95, Weinrich 2007: 328, Engel 2009: 272), which eliminate the neuter candidate, and the default hierarchy, which favours masculine, the gender associated with the noun's form, over feminine, the gender associated with the noun's meaning, contra both the semantics » form approach (1.4.1) and the morphology » semantics » phonology approach (1.4.3).

(6)

	GENDER FEATURES		DEFAULT HIERARCHY		
			*N	*F	*M
<i>Seniorengymnastik</i> 'senior gymnastics'	$-IK \rightarrow *M, *N$	$GAMES/SPORTS \rightarrow *M, *F$			
<i>der Seniorengymnastik</i>	*	*!			*
<i>die Seniorengymnastik</i>		*		*	
<i>das Seniorengymnastik</i>	*		*!		

OGAT is also able to account for the genders of nouns in the sample which have two conflicting features of the same type. As stated in 1.4.1 and 1.4.5.2, many of the other approaches to German gender assignment that have been proposed do not account for such cases. For example, OGAT correctly assigns feminine to the noun *Gedunsenheit* as a result of the GENDER FEATURES constraints $GE \rightarrow *M, *F$ (e.g. Mills 1986: 30, Hickey 2000: 631, Menzel 2004: 66, Rice 2006: 1396, Steinmetz 2006: 1424) and $-HEIT/-(IG)KEIT \rightarrow *M, *N$ (e.g. Zubin & Köpcke 1984: 44, Mills 1986: 30, Flämig 1991: 453, Zifonun et al. 1997: 32, Weinrich 2007: 327, Engel 2009: 281, Eisenberg 2013: 133) eliminating the masculine candidate, and the default hierarchy eliminating the neuter candidate, as shown in (7).

(7)

	GENDER FEATURES		DEFAULT HIERARCHY		
			*N	*F	*M
<i>Gedunsenheit</i> 'swolleness'	$-HEIT/-(IG)KEIT \rightarrow *M, *N$	$GE \rightarrow *M, *F$			
<i>der Gedunsenheit</i>	*	*!			*
<i>die Gedunsenheit</i>		*		*	
<i>das Gedunsenheit</i>	*		*!		

Finally, OGAT is also able to account for the genders of a number of nouns in the sample to which no known, valid GENDER FEATURES constraints apply. As stated in 1.4.5.2, many of the other models of German gender assignment that have been proposed do not provide a strategy for gender assignment in such cases. For instance, OGAT correctly assigns masculine to nouns in the sample such as *Manteltarif*, *Areopag*, *Glockenmantel* and *Dipteros*, despite these not being subject to any known, valid GENDER FEATURES constraints, due to the presence of the default hierarchy, in which *M is the lowest-ranked constraint (as explained in 1.4.5.1). These, and many of the above cases therefore provide support for the default hierarchy *N»*F»*M proposed for German by Steinmetz (1985, 1986, 2001, 2006) and Rice (2006). However, on the basis of the discussion in 1.4.5.3, more thorough testing of this hierarchy with regard to its effectiveness in accounting for the entire sample will take place in 3.4.3. The tableau for *Manteltarif* is shown in (8).

(8)

	GENDER FEATURES	DEFAULT HIERARCHY		
		*N	*F	*M
<i>☞ der Manteltarif</i>				*
<i>die Manteltarif</i>			*!	
<i>das Manteltarif</i>		*!		

In summary, OGAT is able to account for the genders of a wide range of nouns from the sample. These include: nouns with a single formal or semantic feature which is relevant to gender assignment; nouns to which multiple, non-conflicting GENDER FEATURES constraints apply; various monosyllables; nouns whose meaning and form conflict in terms of gender assignment; nouns which have multiple conflicting features of the same type; and nouns to which no known, valid GENDER FEATURES constraints apply. Since many of the nouns in the sample can be accounted for in one of these ways, OGAT, even when using only independently-verified GENDER FEATURES constraints, continues to appear a promising theory of German gender assignment.

However, in order to determine whether OGAT can account for the entire noun sample, a number of further specific cases need to be discussed in more detail. Firstly, there are some issues surrounding particular constraints, meaning that, until these are resolved, it cannot be determined whether OGAT can account for the nouns to which these constraints (potentially) apply. These are the constraints of which, according to Enger's (2009) criteria, two equally plausible versions exist (see 2.9), and a small number of constraints whose relevance to certain nouns in the sample is not entirely clear. This will be discussed in 3.2.

Secondly, there are particular subsets of nouns in the sample which need to be considered more closely in order to determine whether OGAT in its current state is able to account for them, or whether

any adjustments to the theory need to be made. These are: loanwords, nouns with multiple meanings or forms, complex nouns and the subset of 55 nouns, which includes proper nouns and MGNs (see 2.2). These will be considered in 3.3 and some adjustments to OGAT will be proposed on the basis of the discussion.

Thirdly, in 3.4, a number of remaining issues need to be dealt with, including the proposal of a small number of new constraints which are necessary to account for the sample, some proposed further adjustments to OGAT on the basis of some individual nouns, and the testing of alternative rankings within the default markedness hierarchy (see 1.4.5.3).

3.5 will then look at the overall effectiveness of OGAT in accounting for the entire noun sample, and the effectiveness of the original version of OGAT as proposed by Rice (2006) will be compared to that of the version which incorporates all of the adjustments proposed in 3.2 – 3.4, which will be known as OGAT II. This section will also consider the nouns for which neither version can account, the effectiveness of OGAT II in comparison to that of the other approaches outlined in 1.4, and ways in which the theory might be economised. A summary of the findings will be given in 3.6. The predictive powers of OGAT II will then be tested further in Chapter 4 via an experiment with native speakers involving pseudo-nouns.

Specific examples from the noun sample will be used throughout the text, for which in many cases individual OGAT tableaux will be drawn. For purposes of concision, however, a full OGAT tableau will not be drawn for each noun in the sample. A summary table listing each of the nouns in the sample together with the constraints which are relevant to each noun and the predictions of OGAT (and OGAT II) can be found in Appendix F.

3.2 Constraint-related issues

There are a small number of constraints in the constraint set with which certain issues exist, such that these must be dealt with in order to establish whether OGAT can successfully predict the genders of the nouns which are (potentially) subject to these constraints. These are: a small set of constraints with two versions, to be discussed in 3.2.1, and a few constraints whose application to the sample is not straightforward, discussed in 3.2.2.

3.2.1 Constraints with multiple versions

As established in 2.9, there are a number of constraints which, according to Enger's (2009) criteria, have two equally plausible versions, e.g. $-EUR/ÖR \rightarrow *F, *N$ (e.g. Flämig 1991: 453, Hoberg 2004: 86, Weinrich 2007: 326) and $-EUR/ÖR \rightarrow *N$ (e.g. Engel 2009: 279). Since the two versions cannot be

differentiated between by means of the criteria, the more plausible version of the constraint must be established by testing each one on the noun sample to see which accounts best for the data. These constraints are listed below.

CVCC→*F/CVCC→*N	-[I]C MONO→*F/-[I]C MONO→*N	MAMMALS→*F/MAMMALS→*N
-EUR/ÖR→*F,*N/-EUR/ÖR→*N	-LT MONO→*F/-LT MONO→*N	MONOSYLL.→*F/MONOSYLL.→*N
-IT→*F,*N/-IT→*F	-FT/-CHT→*M/-FT/-CHT→*N	ORGANS→*M/ORGANS→*N
-IV→*F/-IV→*M,*F	BIRDS→*F/BIRDS→*N	PICTURES→*F/PICTURES→*M
-SAL→*M/-SAL→*M,*F	DISEASES→*M/DISEASES→*N	STICK→*F/STICK→*N
-TUM→*M,*F/-TUM→*F	DOM.ANIMALS→*F/DOM.ANIMALS→*M	UNITS OF TIME→*F/UNITS OF TIME→*N

In most cases, the more plausible variant can indeed be identified by testing each version on the sample. For instance, on the basis of all nouns in the sample ending in [I] followed by a consonant, such as *das Volt*, *das Bild* and *das Gold*,² it is clear that the variant $-[I]C\ MONO \rightarrow *F$ is more plausible than $-[I]C\ MONO \rightarrow *N$. This is illustrated for *Volt* in (9a) and (9b).³

(9a)

Volt	GENDER FEATURES			DEFAULT HIERARCHY		
	SCI. UNITS→*M,*F	CVCC→*F	-[I]C MONO→*N	*N	*F	*M
☞ <i>der Volt</i>	*					*
<i>die Volt</i>	*	*!			*	
<i>das Volt</i>			*	*!		

(9b)

Volt	GENDER FEATURES			DEFAULT HIERARCHY		
	SCI. UNITS→*M,*F	CVCC→*F	-[I]C MONO→*F	*N	*F	*M
<i>der Volt</i>	*!					*
<i>die Volt</i>	*	*	*!		*	
☞ <i>das Volt</i>				*		

For four of the constraints with two versions, however, the more plausible variant cannot be determined on the basis of the noun sample. This is because OGAT produces identical outcomes with both versions of the constraint. For instance, OGAT is able to correctly predict the masculine gender of all nouns in the sample in *-it* – *Bronzit*, *Kalamit*, *Metropolit*, *Siderit*, *Eremit* and *Satellit* – with both the variant $-IT \rightarrow *F$ (e.g. Durrell 2011: 8) and $-IT \rightarrow *F,*N$ (e.g. Bopp 2000-2018), as illustrated by (10a) and (10b).⁴ In such cases, both variants appear in the summary table in Appendix F in brackets.

² *Volt*, *Bild* and *Gold* as the heads of *Kilovolt*, *Diaphanbild* and *Rauschgold* respectively – see 3.3.3.

³ Incidentally, these tableaux also make a case for the variant $CVCC \rightarrow *F$ over $CVCC \rightarrow *N$. $-LT\ MONO \rightarrow *F$ has not been included in the tableaux as this constraint is already subsumed by $-[I]C\ MONO \rightarrow *F$ – see 3.5.4.

⁴ On the basis of these examples it could also be argued that, for reasons of economy, this constraint can be removed from the constraint set entirely, since neither version is actually necessary in order for OGAT to make the correct prediction – see 3.5.4.

(10a)	<i>Bronzit</i> 'bronzite'	GENDER FEATURES		DEFAULT HIERARCHY		
		ROCKS/MINERALS →*F,*N	-IT→*F,*N	*N	*F	*M
	☞ <i>der Bronzit</i>					*
	<i>die Bronzit</i>	*	*!		*	
	<i>das Bronzit</i>	*	*!	*		

(10b)	<i>Bronzit</i> 'bronzite'	GENDER FEATURES		DEFAULT HIERARCHY		
		ROCKS/MINERALS →*F,*N	-IT→*F	*N	*F	*M
	☞ <i>der Bronzit</i>					*
	<i>die Bronzit</i>	*	*!		*	
	<i>das Bronzit</i>	*!		*		

Finally, there is one constraint of which neither variant is able to account for the sample data, meaning that the nouns in the sample provide no evidence for the presence of either version in the constraint set, despite its plausibility as determined by Enger's (2009) criteria. This is: MONOSYLLABLES→*F/MONOSYLLABLES→*N. With either variant of the constraint, OGAT becomes unable to account for a large number of monosyllables in the sample, including: *Angst, Recht, Volt, Gold* etc.

Moreover, the constraint is in fact unnecessary and undesirable from an economy point of view (see 3.5.4), since there already exist numerous, more specific constraints in the set which apply only to monosyllables, e.g. phonological constraints relating to the structure, such as VC MONO→*M,*F; VCCC MONO→*M; CV MONO→*N. A more general constraint is therefore not required, and as shown above, hinders OGAT from producing the correct result. There is, of course, the possibility that the constraint is ranked lower than other GENDER FEATURES constraints, but this, in many cases, also does not lead OGAT to produce the correct results. It is therefore concluded that neither version of the constraint is likely to be part of the constraint set.

A list of the variants that have been deemed the most plausible as a result of the above findings is given below. With these variants, OGAT is able to account for a high proportion of nouns in the sample to which these constraints apply, specifically 67/79 (85%). This increases to 73/79 (92%) if the various adjustments to the theory proposed in 3.3 and 3.4 are incorporated into the model (see 3.5.2 for the exceptions).

CVCC→*F	-FT/-CHT→*M	UNITS OF TIME→*N
-IV→*M,*F	BIRDS→*F	(DISEASES→*M/DISEASES→*N)
-SAL→*M	DOM.ANIMALS→*F	(ORGANS→*M/ORGANS→*N)
-TUM→*M,*F	MAMMALS→*F	(-EUR/ÖR→*F,*N/-EUR/ÖR→*N)
-[I]C MONO→*F	PICTURES→*M	(-IT→*F,*N/-IT→*F)
-LT MONO→*F	STICK→*N	

3.2.2 Constraints whose application is problematic

The second subset of constraints to be discussed are those whose application to the sample proves problematic in some way. These are: DEVERBAL STEM \rightarrow *F,*N; SUPERORDINATES \rightarrow *M,*F; DEVERBAL -T \rightarrow *M,*N; and -MA \rightarrow *M,*F. Despite all four of these constraints meeting the necessary criteria in Chapter 2, it is not entirely clear exactly to which nouns in the sample native speakers will apply them. For DEVERBAL STEM \rightarrow *F,*N and SUPERORDINATES \rightarrow *M,*F, this is related to the clarity of what the constraint covers (criterion 4, Section 2.6), and for DEVERBAL -T \rightarrow *M,*N and -MA \rightarrow *M,*F, this is related to the productivity of the feature involved (criterion 6, Section 2.8). Consequently, in order to determine whether OGAT can account for the genders of the nouns to which these constraints might potentially apply, these issues must be explored further. Each of the four constraints will be discussed individually below.

3.2.2.1 DEVERBAL STEM \rightarrow *F,*N

The first constraint which is problematic in terms of its application is DEVERBAL STEM \rightarrow *F,*N, which dictates that nouns derived from bare verb stems, e.g. *der Versuch* (from *versuchen*), including ablauted verb stems, e.g. *der Gang* (from *gehen*), cannot be feminine or neuter (e.g. Flämig 1991: 453, Hickey 2000: 643, Hoberg 2004: 90, Chan 2005: 55, Fleischer & Barz 2012: 267f.). As stated in 1.3.1.2, nouns of this type are generally considered to be the product of implicit/zero derivation or deverbal conversion (for an overview of the debate as to which of these terms is more appropriate, see e.g. Kastovsky 2005: 31-36).

Irrespective of how the process is categorised, it is the directionality of this process which poses a problem for the application of the DEVERBAL STEM \rightarrow *F,*N constraint.⁵ The constraint can, of course, only be applied to nouns which are perceived by native speakers to have been derived from verbs. Nouns which have a corresponding verb, but which are not perceived to have been derived from that verb (the verb instead having been derived from the noun), are not subject to the constraint. The problem is that it is not always clear in which direction the implicit derivation/zero derivation/conversion (henceforth simply derivation) is perceived to have occurred.

For nouns with a stem vowel which differs to that of its corresponding verb, e.g. *der Bruch – brechen*, it is widely accepted that the noun is a deverbal derivative and, given the formal indication, likely to be perceived by native speakers as such (e.g. Naumann & Vogel 2000: 932, Eisenberg 2013: 295,

⁵ Assuming there exists a directional relationship at all, cf. Lieber (1980) (see discussion in e.g. Balteiro 2007: 62-65, 83-88).

Fleischer & Barz 2012: 89f. etc.). However, in cases where the noun and corresponding verb stem are identical, e.g. *der Fall* – *fallen*, it can be extremely difficult to determine whether the verb is perceived to be the source of the noun (and therefore subject to the constraint $\text{DEVERBAL STEM} \rightarrow *F, *N$) or whether the noun is in fact perceived to be the source of the verb (and therefore not subject to the constraint).

There are around 40 nouns in the sample to which the constraint $\text{DEVERBAL STEM} \rightarrow *F, *N$ might plausibly apply. 14 of these, listed below, have a stem vowel which is different to that of their corresponding verb, indicating that these are likely to be recognised as deverbal derivatives and therefore subject to the constraint $\text{DEVERBAL STEM} \rightarrow *F, *N$.

<i>der Ausdrusch (ausdreschen)</i>	<i>der Geschäftsschluss (schließen)</i>	<i>der Verdruss (verdrießen)</i>
<i>der Besatz (besetzen)</i>	<i>der Hochgenuss (genießen)</i>	<i>der Wartestand (stehen)</i>
<i>der Dreiklang (klingen)</i>	<i>der Out-of-Area-Einsatz (einsetzen)</i>	<i>der Weitschuss (schießen)</i>
<i>der Einriss (einreißen)</i>	<i>der Strampelanzug (anziehen)</i>	<i>der Zusammenbruch (brechen)</i>
<i>der Fladerschnitt (schneiden)</i>	<i>das Uhrband (binden)</i>	

For 26 of the nouns, however, the direction of derivation is not so clear, since the noun and the corresponding verb stem are identical. Consequently, it is uncertain whether the constraint $\text{DEVERBAL STEM} \rightarrow *F, *N$ applies to these nouns or not.

<i>der Anreiz (anreizen)</i>	<i>der Ehe Streit (streiten)</i>	<i>der Reinfall (fallen)</i>
<i>die Antwort (antworten)</i>	<i>der Erweiterungsbau (bauen)</i>	<i>die Schwarzarbeit (arbeiten)</i>
<i>der Audiostream (streamen)</i>	<i>der Feldverweis (verweisen)</i>	<i>der Sinnenrausch (rauschen)</i>
<i>der Auftrag (auftragen)</i>	<i>der Fluchtversuch (versuchen)</i>	<i>der Steuerfreibetrag (betragen)</i>
<i>der Avistawechsel (wechseln)</i>	<i>das Gebärdenspiel (spielen)</i>	<i>der Strahl (strahlen)</i>
<i>der Bürgersteig (steigen)</i>	<i>der Hops (hopsen)</i>	<i>der Test (testen)</i>
<i>der Containerverkehr (verkehren)</i>	<i>der Kauerstart (starten)</i>	<i>die Umschau (umschauen)</i>
<i>der Detailhandel (handeln)</i>	<i>der Nachschlag (nachschießen)</i>	<i>der Verkauf (verkaufen)</i>
<i>der Dienstgebrauch (gebrauchen)</i>	<i>die Regel (regeln)</i>	

Fox (2005: 139) states that, in such cases, the verb is normally considered to be the source of the noun, however he acknowledges that with pairs such as *teilen* ‘to share’ – *Teil* ‘piece’, it seems more likely that the noun is the source of the verb. Since the noun and verb stem are identical, the direction of derivation cannot, of course, be determined on the basis of word structure itself. We must, then, seek other evidence.

One possibility is to use diachronic data to determine the direction of derivation by observing whether the verb or noun was attested first (an approach favoured by e.g. Štekauer 1996: 133). However, Marchand (1963, 1964, 1969), argues that it is virtually impossible to determine the direction of this relationship using diachronic methods. He maintains that diachronic data prove ineffective when the first attestations of the verb and noun are from around the same date, and for words which date

relatively far back (e.g. to Old High German), the first attestations of words often cannot be determined due to the scarcity of records (Marchand 1963: 226). Moreover, the first attestation of a word in the textual record need not have any bearing on its first attestation in speech.

Furthermore, Plag (2003: 136) observes that there are word pairs for which the diachronic evidence is at odds with the intuition of present-day native speakers, which is what is of interest to this study. He argues that most native English speakers would, for example, claim that the verb *to crowd* derives from the noun *crowd*. However, the OED states that the verb was attested first. Similarly, Marchand (1963: 229f.) notes that the noun *moan* is recorded much earlier than the verb *to moan* (ca. 1225 vs. ca. 1548), yet this is commonly analysed synchronically as a deverbal derivative. It seems, therefore, that the directionality problem cannot be solved using diachronic data, since the original direction of derivation may have been overwritten by ‘complex semantic changes’ (Plag 2003: 136).

Given the inability of diachronic methods to model contemporary native speaker competence, many linguists have instead turned to synchronic evidence for the solution. A variety of criteria have been proposed which seek to determine the direction of derivation through the investigation of various synchronic factors (see Marchand 1963, 1964; Olsen 1986: 122f.; Balteiro 2007; Umbreit 2010, 2015; Fleischer & Barz 2012: 268f.). Three of the main criteria will be discussed below.

The first and perhaps most well-known criterion is that of semantic dependence. Marchand (1964: 12) states: ‘the word that for its analysis is dependent on the content of the other pair member is necessarily the derivative’. He illustrates this using the English word pair *saw_N* and *saw_V*. *Saw_N*, he argues, can be adequately defined without reference to the act of sawing, e.g. ‘tool with a thin, serrated blade’. However, the definition of *saw_V* necessarily refers to *saw_N*: ‘to cut with a saw’. It can therefore be deduced that *saw_V* is derived from *saw_N* (Marchand 1964: 12).

This criterion also seems promising for German. For instance, Fleischer & Barz (2012: 269) reason that the verb *fischen* ‘to fish’ is motivated by the noun *der Fisch* ‘fish’, given that it is necessary to refer to the noun when defining the verb, whilst the reverse is not true. *Fischen* can be defined as ‘*Fische fangen* [to catch fish]’, yet defining *Fisch* does not require reference to being caught. Similarly, with the word pair *Schau* ‘show’ – *schauen* ‘to look’, Fleischer & Barz (2012: 269) argue that it is clear that the meaning of *schauen* is not ‘*sich mit einer Schau beschäftigen* [to occupy oneself with a show]’, and so the verb must be seen as the base and noun as a deverbal derivative.

However, the semantic dependence criterion has been criticised for its lack of objectivity. Both Štekauer (1996: 128) and Umbreit (2010: 310) argue that application of the criterion relies heavily on the personal interpretation of the words in question. Štekauer (1996: 128) maintains that ‘we can

adjust the definition of semantically related words in accordance with our intentions'. Regarding the *saw_N* – *saw_V* example given by Marchand (1964: 12), Štekauer (1996: 128) claims that the case can easily be made for the opposite direction of derivation (i.e. *saw_V* as the base) by defining the verb as 'to cut with a toothed instrument' and the noun as 'instrument for sawing'. However, this reasoning is hard to apply in cases such as *Fisch* – *fischen*, as it seems highly unlikely that *Fisch* could ever be defined as 'etwas, was gefischt wird [something which is fished]'.

Despite the potential subjectivity of the criterion, Umbreit (2010: 310) acknowledges that it could still be of use if the preference for one interpretation over another is universal. While this might appear difficult to establish in practical terms, a solution could consist of analysing the definitions of each pair member in an independent source. Accordingly, this has been done for the nouns in the sample whose corresponding verb stem is identical using definitions from the Duden online dictionary.

Noun from sample and corresponding verb	Reference made to the other pair member?		Deverbal?
	Noun	Verb	
<i>Anreiz (anreizen)</i>	no	yes	x
<i>Antwort (antworten)</i>	no	yes	x
<i>Arbeit (arbeiten)</i>	yes	yes	-
<i>Audiostream (streamen)</i>	no	yes	x
<i>Auftrag (auftragen)</i>	no	yes	x
<i>Bau (bauen)</i>	yes	yes	-
<i>Betrag (betragen)</i>	no	no	-
<i>Fall (fallen)</i>	yes	no	✓
<i>Gebrauch (gebrauchen)</i>	yes	no	✓
<i>Handel (handeln)</i>	yes	yes	-
<i>Hops (hopsen)</i>	yes	no	✓
<i>Nachschlag (nachschiagen)</i>	no	no	-
<i>Rausch (rauschen)</i>	no	no	-
<i>Regel (regeln)</i>	no	yes	x
<i>Spiel (spielen)</i>	yes	yes	-
<i>Start (starten)</i>	yes	no	✓
<i>Steig (steigen)</i>	no	no	-
<i>Strahl (strahlen)</i>	no	yes	x
<i>Streit (streiten)</i>	no	yes	x
<i>Test (testen)</i>	no	yes	x
<i>Umschau (umschauen)</i>	yes	no	✓
<i>Verkauf (verkaufen)</i>	yes	no	✓
<i>Verkehr (verkehren)</i>	no	no	-
<i>Versuch (versuchen)</i>	yes	no	✓
<i>Verweis (verweisen)</i>	no	no	-
<i>Wechsel (wechseln)</i>	yes	no	✓

Table 1: Direction of derivation according to the semantic dependence criterion

The results show some clear cases of verb definitions making reference to the corresponding noun, suggesting that the verb is perceived as a denominal derivative. For instance, the definition of *antworten* 'to answer' is listed as 'Antwort, Auskunft geben [give an answer/information]' and *streiten* 'to argue' is defined as 'mit jemandem Streit haben, in Streit geraten [to have/get into an argument with someone]'. Conversely, there are cases of the noun definition making reference to the verb, indicating that the verb is perceived as the derivational base. For example, the definition of *Gebrauch* 'use' is listed as 'das Gebrauchen [use]', and the definition of *Versuch* 'attempt' contains 'Handlung, mit der etwas versucht wird [act by which something is attempted]'.

In some cases, the criterion cannot be used to determine the direction of derivation, since both members of the word pair are defined without making reference to the other pair member, e.g. *Betrag* – *betragen*, *Steig* – *steigen*, *Verkehr* – *verkehren*. There are also a few cases in which reference to the other pair member is made in the definitions of both the noun and the verb, namely: *Bau* – *bauen*, *Handel* – *handeln*, *Arbeit* – *arbeiten* and *Spiel* – *spielen*. The semantic dependence criterion does not prove useful in these cases.

Another criterion which has been suggested for determining the direction of derivation concerns frequency of usage. It is claimed that the word with lower frequency is perceived to be derived from the word with higher frequency (e.g. Plag 2003). This criterion is in line with the tendency that derived units are often more semantically complex than their base, i.e. often have a more specific meaning (Marchand 1964: 13, Iacobini 2000: 866), so it follows that derivatives fit into fewer contexts and therefore have a lower frequency (Marchand 1964: 13, Plag 2003: 111).

However, methodologically, frequency can be extremely difficult to determine (e.g. Fleischer & Barz 2012: 269). Umbreit (2010: 308f.), for instance, highlights that frequency calculations vary greatly depending on the sources examined, given that certain sources may favour the use of the noun over the verb, or vice versa, according to the text type. Indeed, it is widely reported that newspapers, for instance, tend to prefer the use of nouns over verbs (e.g. Popadić 1971; Jesenšek 1995: 175, 179; Lüger 1995: 12, 25 etc.). Umbreit (2010: 309) also maintains that this criterion is of limited use in cases where there is not a significant difference between the frequency of the verb and corresponding noun.

In order to overcome these potential difficulties, the frequency of the 26 noun/verb pairs in question has been assessed using the DWDS-Kernkorpus 21 (Digitales Wörterbuch der deutschen Sprache), which contains a total of 15,462,297 tokens taken from a range of texts (fiction, non-fiction, scientific texts and newspapers) from the period 2000-2010. Calculating frequency across such a broad range

of text types should prevent the problem raised by Umbreit (2010: 308f.) that certain text types may have stylistic preferences for using nouns over verbs or vice versa.

Noun from sample and corresponding verb	Frequency per million tokens (absolute figure)		Deverbal?
	Noun	Verb	
<i>Anreiz (anreizen)</i>	12.1 (4,309)	0.3 (40)	x
<i>Antwort (antworten)</i>	113.7 (35,564)	52.4 (17,533)	x
<i>Arbeit (arbeiten)</i>	309.9 (103,447)	246.0 (100,604)	x
<i>Audiostream (streamen)</i>	0.93 (289)	0.17 (30)	x
<i>Auftrag (auftragen)</i>	78.2 (26,397)	3.5 (1,274)	x
<i>Bau (bauen)</i>	90.6 (34,170)	132.3 (47,297)	✓
<i>Betrag (betragen)</i>	25.7 (8,177)	62.2 (18,555)	✓
<i>Fall (fallen)</i>	345.8 (121,601)	162.4 (56,450)	x
<i>Gebrauch (gebrauchen)</i>	22.6 (5,571)	31.8 (11,198)	✓
<i>Handel (handeln)</i>	64.5 (15,855)	182.0 (58,102)	✓
<i>Hops (hopsen)</i>	0.4 (541)	0.6 (245)	✓
<i>Nachschlag (nachschießen)</i>	1.0 (286)	2.0 (520)	✓
<i>Rausch (rauschen)</i>	7.3 (2,726)	4.7 (1,962)	x
<i>Regel (regeln)</i>	100.3 (35,801)	29.3 (10,360)	x
<i>Spiel (spielen)</i>	315.0 (101,833)	353.8 (139,965)	✓
<i>Start (starten)</i>	57.7 (18,338)	74.2 (28,746)	✓
<i>Steig (steigen)</i>	0.6 (200)	204.5 (70,697)	✓
<i>Strahl (strahlen)</i>	5.5 (1,654)	13.9 (5,740)	✓
<i>Streit (streiten)</i>	88.1 (35,220)	34.1 (14,168)	x
<i>Test (testen)</i>	32.54 (14,703)	22.94 (10,624)	x
<i>Umschau (umschauen)</i>	0.44 (83)	1.66 (837)	✓
<i>Verkauf (verkaufen)</i>	55.9 (22,083)	131.8 (52,765)	✓
<i>Verkehr (verkehren)</i>	36.4 (10,726)	10.0 (3,175)	x
<i>Versuch (versuchen)</i>	114.8 (29,355)	240.5 (80,428)	✓
<i>Verweis (verweisen)</i>	6.9 (2,949)	60.5 (20,092)	✓
<i>Wechsel (wechseln)</i>	41.39 (12,911)	63.19 (24,542)	✓

Table 2: Direction of derivation according to frequency of usage

The data show that for 11 out of the 26 word pairs, the noun is listed as having a higher frequency than the corresponding verb, which, according to the criterion, suggests that the noun is likely to be perceived as the source of the verb. There are 15 cases in which the verb has a higher frequency than the corresponding noun, indicating that the verb is likely to be perceived as the derivational base and the noun as a deverbal derivative. There appear to be no instances of the verb and noun having the same or a highly similar frequency, rendering Umbreit (2010: 309)'s concern about there not being a difference between the frequency of the verb and noun irrelevant in this case. In most cases, the results of this criterion correspond to that of the semantic dependence criterion. There are just two cases which do not: *Fall – fallen* and *Strahl – strahlen*.

The final criterion to be discussed in this section is one concerning the structural properties of the noun. There are certain prefixes in German which are limited to the formation of verbs, e.g. *be-*, *ent-*, *er-*, *ver-* and *zer-*. Fleischer & Barz (2012: 268) state that any noun which has one of these prefixes must therefore be a deverbal derivative. Also belonging to this category are nouns containing verbal particles, such as *ab-*, *an-*, *auf-*, *ein-*, *nach-*, *um-* and *vor-*.

There are two main limitations to this criterion. The first is that, as previously mentioned, it is possible that the direction of derivation for certain word pairs may have come to be thought of differently by present-day native speakers due to semantic changes. The second limitation is that the criterion can, of course, only determine the direction of derivation for nouns and verbs which contain these prefixes or particles. However, of the 26 word pairs in question, it is able to make a judgement for over a third of these cases, shown below.

Noun from sample	Prefix/particle	Deverbal?
<i>Anreiz (anreizen)</i>	<i>an-</i>	✓
<i>Auftrag (auftragen)</i>	<i>auf-</i>	✓
<i>Betrag (betragen)</i>	<i>be-</i>	✓
<i>Nachschlag (nachschiagen)</i>	<i>nach-</i>	✓
<i>Umschau (umschauen)</i>	<i>um-</i>	✓
<i>Verkauf (verkaufen)</i>	<i>ver-</i>	✓
<i>Verkehr (verkehren)</i>	<i>ver-</i>	✓
<i>Versuch (versuchen)</i>	<i>ver-</i>	✓
<i>Verweis (verweisen)</i>	<i>ver-</i>	✓
<i>Antwort (antworten)</i>	<i>(ant-)*</i>	-

According to the criterion, nine of the nouns in question are likely to be perceived as deverbal derivatives, since they contain either a verbal prefix or particle. These results correspond with the results from the other two criteria for six of these cases.

It should be noted that the case of *Antwort* – *antworten* is somewhat unusual with regard to this criterion, given that the original verbal prefix, *ant-*, is now virtually obsolete. Kluge (2002: 16) observes that it remains visible in just two cases: *Antlitz* and *Antwort*. It is therefore highly likely that present-day native speakers analyse *Antwort* as a simplex noun and that the original prefix *ant-* thus has no impact on synchronic perceptions of the direction of derivation. Additionally, the – perhaps circular – argument might be made that the fact that *Antwort*, a feminine noun, does not adopt the neuter gender of *Wort* is likely to reinforce the idea that it is not a prefixed noun.

Summary

Despite their potential shortcomings, the three criteria of semantic dependence, frequency of usage and presence of verbal prefix/particle are able to provide an indication of the perceived direction of derivation for many of the noun-verb pairs in question. A summary of the results is given below.

Noun from sample and corresponding verb	Deverbal?			
	Semantic dependence	Frequency	Prefix/particle	Conclusion
<i>Verkauf (verkaufen)</i>	✓	✓	✓	yes
<i>Fluchtversuch (versuchen)</i>	✓	✓	✓	yes
<i>Umschau (umschauen)</i>	✓	✓	✓	yes
<i>Hops (hopsen)</i>	✓	✓	-	yes
<i>Kauerstart (starten)</i>	✓	✓	-	yes
<i>Dienstgebrauch (gebrauchen)</i>	✓	✓	-	yes
<i>Avistawechsel (wechseln)</i>	✓	✓	-	yes
<i>Steuerfreibetrag (betragen)</i>	-	✓	✓	yes
<i>Nachschlag (nachschiessen)</i>	-	✓	✓	yes
<i>Feldverweis (verweisen)</i>	-	✓	✓	yes
<i>Erweiterungsbau (bauen)</i>	-	✓	-	yes
<i>Detailhandel (handeln)</i>	-	✓	-	yes
<i>Bürgersteig (steigen)</i>	-	✓	-	yes
<i>Gebärdenspiel (spielen)</i>	-	✓	-	yes
<i>Ehrestreit (streiten)</i>	x	x	-	no
<i>Audiostream (streamen)</i>	x	x	-	no
<i>Test (testen)</i>	x	x	-	no
<i>Regel (regeln)</i>	x	x	-	no
<i>Antwort (antworten)</i>	x	x	-	no
<i>Sinnenrausch (rauschen)</i>	-	x	-	no
<i>Schwarzarbeit (arbeiten)</i>	-	x	-	no
<i>Anreiz (anreizen)</i>	x	x	✓	no
<i>Auftrag (auftragen)</i>	x	x	✓	no
<i>Reinfall (fallen)</i>	✓	x	-	?
<i>Strahl (strahlen)</i>	x	✓	-	?
<i>Containerverkehr (verkehren)</i>	-	x	✓	?

The results indicate that 14 of the problematic nouns in the sample are likely to be perceived as deverbal derivatives and nine are not. In three cases (*Verkauf*, *Versuch* and *Umschau*), all three criteria are unanimous in their verdict, in 14 cases, the conclusion has been drawn based on two out of the three criteria, and in six cases, only the second criterion could provide an indication of the perceived direction of derivation.

There are three nouns in the set for which a judgement is unable to be made based on the results, since only two criteria were able to produce a result and these results conflict. Fortunately for our analysis, neither the presence nor absence of this constraint affects OGAT's ability to predict the

correct gender for these three nouns, as shown below. In such cases, the constraint DEVERBAL STEM→*F,*N will be listed in parentheses in the table in Appendix F.

Noun	Constraints	OGAT prediction
<i>Fall</i> (<i>der Reinfeld</i>)	-L MONO→*F; CVC MONO→*F; (DEVERBAL STEM→*F,*N)	✓ (M)
<i>Strahl</i> (<i>Korpuskularstrahlen</i>)	CCCVC MONO→*F,*N; STICK→*N; [[]C- MONO→*F,*N; V _{LONG} (C) MONO→*F; -L MONO→*F; (DEVERBAL STEM→*F,*N)	✓ (M)
<i>Verkehr</i> (<i>der Containerverkehr</i>)	(DEVERBAL STEM→*F,*N)	✓ (M)

Now that it has been established to which nouns in the sample native speakers are likely to apply the constraint DEVERBAL STEM→*F,*N, it is possible to assess OGAT’s ability to account for the genders of these nouns. The analysis demonstrates that OGAT yields the correct prediction in a high proportion of cases, specifically 92% (see 3.5.2 for exceptions). Two such cases are illustrated below.

(11)	<i>Gebrauch</i> ‘use’	GENDER FEATURES		DEFAULT HIERARCHY		
		GE-→*M,*F	DEVERBAL STEM→*F,*N	*N	*F	*M
	<i>☞ der Gebrauch</i>	*				*
	<i>die Gebrauch</i>	*	*!		*	
	<i>das Gebrauch</i>		*	*!		

(12)	<i>Klang</i> ‘sound’	GENDER FEATURES				DEFAULT HIERARCHY		
		DEVERBAL STEM→*F,*N	-ANG→*F,*N	CCVC MONO→*F,*N	-N(c)(c)→*F,*N	*N	*F	*M
	<i>☞ der Klang</i>						*	
	<i>die Klang</i>	*	*	*	*!		*	
	<i>das Klang</i>	*	*	*	*!	*		

3.2.2.2 SUPERORDINATES→*M,*F

The second constraint whose application proves problematic is SUP.→*M,*F “nouns denoting superordinate categories cannot be masculine or feminine”. While this principle is attested by numerous linguists, including Zubin & Köpcke (1981, 1986), Mills (1986: 27), Steinmetz (1986: 192, 2006: 1424), Köpcke & Zubin (1997: 35-38), Eisenberg (1999: 156) and Rice (2006: 1398), it is not always clear to which nouns native speakers are likely to apply the constraint.

Superordinates are commonly thought of as nouns denoting categories or as “umbrella” terms for a class of objects, such as *Obst* ‘fruit’, *Gemüse* ‘vegetable’ and *Fahrzeug* ‘vehicle’. However, this definition is relatively vague, meaning that the nouns which are subject to the constraint SUP.→*M,*F cannot always be easily identified. For instance, *Obst* and *Gemüse* may typically be thought of as

superordinates, yet these terms themselves are also hyponyms of the more general category *das Essen* 'food'. Similarly, certain types of fruit and vegetables could also be considered umbrella terms, e.g. *Zwiebel* 'onion' is itself the head of a category which includes hyponyms such as *Frühlingszwiebel* 'spring onion' and *Schalotte* 'shallot'.

In the noun sample, there are numerous nouns which could be considered superordinates as they denote categories which contain various hyponyms. These include: *Haus* 'house', *Baum* 'tree', *Spiel* 'game' and *Vogel* 'bird'. However, their status as superordinates is not entirely clear, given that some of these could also be classed as hyponyms of other categories (e.g. *Baum* < *Gewächs* 'plant-life', *Vogel* < *Tier* 'animal').

In order to gain a clearer understanding of superordinates, we must first consider the hierarchical structure within which objects are organised, briefly introduced in 1.3.3.3. Rosch (1977) differentiates between three levels of abstraction: superordinate > basic level > subordinate. The higher up the hierarchy one goes, the more inclusive the terms are.

Superordinate	Basic Level	Subordinate
<i>das Instrument</i>	<i>die Gitarre</i> 'guitar' <i>das Klavier</i> 'piano' <i>die Trommel</i> 'drum'	<i>die Konzertgitarre</i> 'concert guitar' <i>der Flügel</i> 'grand piano' <i>die Basstrommel</i> 'bass drum'
<i>das Obst</i> 'fruit'	<i>der Apfel</i> 'apple' <i>der Pfirsich</i> 'peach' <i>die Traube</i> 'grape'	<i>der Granny-Smith</i> <i>der kalifornische Pfirsich</i> 'California peach' <i>die kernlose Traube</i> 'seedless grape'
<i>das Werkzeug</i> 'tool'	<i>der Hammer</i> <i>die Säge</i> 'saw' <i>der Schraubenzieher</i> 'screwdriver'	<i>der Zimmermannshammer</i> 'claw hammer' <i>die Blattsäge</i> 'pad saw' <i>der Kreuzschraubenzieher</i> 'crosstip screwdriver'
<i>das Fahrzeug</i> 'vehicle'	<i>der Wagen</i> 'car' <i>der Bus</i> <i>der Laster</i> 'lorry'	<i>der Sportwagen</i> 'sports car' <i>der Stadtbus</i> 'city bus' <i>der Kipplaster</i> 'dumper truck'
<i>das Tier</i> 'animal'	<i>der Fisch</i> 'fish' <i>der Vogel</i> 'bird'	<i>die Forelle</i> 'trout' <i>der Adler</i> 'eagle'

Table 3: Levels of categorisation (Köpcke & Zubin 1997: 36, originally adapted from Rosch 1977: 32)

The level of abstraction which is key to this hierarchy is the basic level, since it is the basic level around which all other levels are defined (e.g. Rosch et al. 1976, Rosch 1977). That is to say, superordinates are terms on the level of abstraction above the basic level, and subordinates are terms organised on the level below the basic level. It is therefore necessary to determine the basic level in order to determine a noun's hierarchical status.

Ideas concerning the presence of a basic level date back to Brown (1958, 1965), who hints at there being a level of abstraction which is the most culturally salient for speakers: 'While a dime *can* be called a *coin* or *money* or a *1952 dime*, we somehow feel that *dime* is its real name. The other categorizations seem like achievements of the imagination' (Brown 1965: 320). Yet the question of

how exactly basic level terms can be recognised remains. Evidence from psycholinguistics reveals that basic level categories can be recognised through the analysis of four main features: cognition, perception, communication and interaction, each of which will be discussed below.

Firstly, the basic level is the level at which things seem to be cognitively organised. Evidence for this includes the fact that it is basic level terms which are used the most frequently in speakers' descriptions of things (e.g. Berlin et al. 1973, Rosch et al. 1976). In an experiment conducted by Rosch et al. (1976), speakers were shown pictures depicting specific varieties of objects and were asked to describe what they saw. It was found that basic level terms were most commonly used. For instance, when shown a specific variety of apple, the participants tended to label this using the basic level term "apple", rather than the superordinate "fruit" or the subordinate "Granny Smith". This occurred despite the fact that the participants showed awareness of both the superordinate and subordinate terms when questioned. Additionally, when the participants were presented with a picture and then asked whether or not a particular label applied to it, the participants were able to make the quickest decisions when given basic level terms rather than superordinate or subordinate terms. These findings suggest that cognitive organisation occurs primarily on the basic level.

Secondly, the basic level can also be determined by analysing the ways in which objects are perceived. Rosch (1977: 31) claims that the basic level is the level at which there are perceived to be the highest number of attributes which are shared by all members of that category. For example, there are more attributes perceived to be shared by all trains than by all vehicles in general. The basis of this observation lies in ethnobiological research, which demonstrates that in a biological taxonomy, the genus is the level at which organisms have the most attributes in common with each other whilst simultaneously preserving a large number of differences with other classes (e.g. Berlin 1978). Experimental data from Rosch et al. (1976) largely confirms this. Participants were given lists of words from differing levels of abstraction and were asked to list as many attributes of each item as possible. The results showed that participants had listed very few attributes which they perceived to be common to superordinate terms, a significantly larger number for basic level terms, and around the same number for subordinate terms. For instance, participants were able to list only a few attributes for *furniture* (as there is little that all items of furniture have in common), but significantly more for *table* (as there are more features all tables have in common), and then around the same number for *coffee table*.

One of the most interesting findings from this experiment is that the results from the three biological taxonomies tested (*tree*, *fish* and *bird*) were different from the results expected according to the

ethnobiological research by Berlin (1972, 1978) and Berlin et al. (1974) etc. The ethnobiologists claimed that it is the genus of a biological taxonomy which should be equivalent to the basic level, e.g. *oak, birch, trout, salmon*. Rosch et al. (1976), however, demonstrated that this is not the case, and that it is instead terms such as *tree, fish* and *bird* which are on the basic level, psycholinguistically speaking, as shown below.

Superordinate	Basic Level	Subordinate
<i>plant-life (das Gewächs)</i>	<i>tree (der Baum)</i>	<i>birch (die Birke)</i> <i>oak (die Eiche)</i>
<i>animal (das Tier)</i>	<i>fish (der Fisch)</i>	<i>salmon (der Lachs)</i> <i>trout (die Forelle)</i>
	<i>bird (der Vogel)</i>	<i>blackbird (die Amsel)</i> <i>sparrow (der Spatz)</i>

Building on the findings of Rosch et al. (1976), research by Tversky (1985) shows that one of the reasons why so many common attributes can be listed at the basic level is that the basic level is the highest level at which parts of an object can be specified. For example, the basic level term *chair* can be easily described as a sum of its parts: seat, legs and back, yet it is impossible to list parts which are common to all items of the superordinate *furniture*. This is also further justification that terms such as *bird, fish* and *tree* are on the basic level, as parts of these can be readily listed (e.g. wings and beak, gills and fins etc.), whereas it is difficult to do the same for all animals or all plant life.

A further experiment by Rosch et al. (1976) demonstrates that the basic level is also the highest level at which objects have the most similarly perceived overall shapes. In the experiment, participants were given a list of objects (from all three levels of abstraction) and were asked to draw the outline of each of them. Upon comparing the outlines, it was observed that there was very little overlap between objects of the same superordinate category (e.g. different vehicles), but a significantly higher degree of overlap between objects from the same basic level category (e.g. different types of car). There was no significant increase in the degree of overlap between members of the same subordinate category (e.g. different sportscars).

Thirdly, the basic level is also the level which is most relevant in terms of certain aspects of communication. Rosch (1999: 258) observes that the basic level is the first level named and understood by children. That is to say, children tend to produce and comprehend nouns denoting basic level objects before they are able to produce and comprehend nouns denoting objects on the superordinate or subordinate levels. Additionally, and most probably related to this, is the observation that words on the basic level tend to be relatively short. If we take into consideration the finding by Zipf (1935) that there is a negative correlation between word length and word frequency (i.e. the most

frequent terms tend to be short), coupled with the observation that basic level terms tend to be the most frequent (discussed above), then it should follow that basic level terms are shorter and morphologically simpler than superordinate and subordinate terms. Indeed, both Berlin et al. (1973) and Casson (1981) have shown this to be true. Berlin et al. (1973), for example, observe that subordinate level terms commonly consist of a “primary” lexeme plus a modifier, e.g. *coffee table*, whereas basic level terms ordinarily consist just of a “primary” lexeme, e.g. *table*.

Fourthly and finally, the basic level can also be determined by analysing the ways in which people physically interact with objects. In an experiment, Rosch et al. (1976) asked participants to provide a detailed description of the series of movements they performed when physically interacting with different objects. The movements which all participants had listed were added up. The participants were found to have all listed the most motor movements for objects at the basic level. This was to be expected, since there are few movements that humans carry out when interacting with e.g. all furniture, but a considerably greater number of different movements that all humans carry out when interacting with, for instance, a chair (e.g. Rosch 1999).

To summarise, it has been established that in order to determine whether a noun denotes a superordinate or not, the basic level must first be identified. To identify the basic level, a number of questions can be asked:

- If shown a picture of a particular variety of X, what would be the instinctive label that we give it?
- Can we list numerous attributes common to all Xs (or at least significantly more than for a noun at a higher level of abstraction)?
- Can we list individual parts of X?
- If we were to draw outlines of X, would these be similar to one another?
- Are children likely to know this word?
- Is the word relatively short?
- Is the word relatively frequent?
- Can we list multiple ways in which we might interact with the object (or at least more ways than for an object at a higher level of abstraction)?

Once the basic level terms have been identified, the superordinate level term can then be determined by moving one level of abstraction higher.

Zubin & Köpcke (1986) and Köpcke & Zubin (1997: 35-38), who originally highlighted the connection between German nouns at the superordinate level and the neuter gender, add a further method of superordinate recognition for German which concerns the form of the noun. They observe that superordinates in German often consist of compounds whose right-hand element (or head, see 3.3.3) is a noun with a vague meaning such as *Zeug* ‘stuff’, *Mittel* ‘means’ or *Ding* ‘thing’, e.g. *Lebensmittel*

'food' and *Fahrzeug* 'vehicle', or of derived nouns which are prefixed with *Ge-*, e.g. *Getreide* 'crop' and *Getränk* 'drink'.

Additionally, Köpcke & Zubin (1997: 37) observe that it is possible to add further levels of abstraction to the hierarchy, as shown in the table below. This further highlights how central the basic level is to the definition of all levels of abstraction, since the basic level does not change, regardless of how many levels of abstraction are added above or below it.

Superordinate (1)	Superordinate (2)	Basic Level	Subordinate
<i>das Lebensmittel</i> 'food'	<i>das Gemüse</i> 'vegetable'	<i>der Kohl</i> 'cabbage' <i>die Möhre</i> 'carrot'	<i>der Wirsingkohl</i> 'savoy cabbage' <i>die Chantenay-Möhre</i> 'Chantenay carrot'
	<i>das Obst</i> 'fruit'	<i>die Dattel</i> 'date' <i>der Apfel</i> 'apple'	<i>die Medjool Dattel</i> 'medjool date' <i>der Granny-Smith</i>
	<i>das Fleisch</i> 'meat'	<i>das Schweinefleisch</i> 'pork' <i>das Lammfleisch</i> 'lamb'	<i>das iberische Schweinefleisch</i> 'Iberian pork' <i>das Heidschnuckenfleisch</i> 'German heath lamb'

In terms of the sample, then, nouns which native speakers are likely to categorise on the superordinate level, and are therefore subject to the constraint SUP. → *M, *F, can now be more easily identified, since practical measures for recognising basic level terms have been established.

For instance, using the set of questions listed above, it can be reasoned that *Haus* 'house' is not a superordinate but a basic level term. Firstly, if shown a picture of a particular type of house (e.g. detached, terraced), it is highly likely that a speaker would instinctively label it as a house. It is also possible to list numerous attributes of a house, whereas this is slightly harder for the more general term *dwelling*. Additionally, individual parts of a house can be readily listed (e.g. door, windows, stairs). Finally, children are likely to know the word *Haus* and it is relatively short in length. Similar reasoning can be provided for other nouns in the sample such as *Baum* 'tree' and *Wagen* 'car' (see also Zubin & Köpcke 1986).

Conversely, it is clear from the set of questions that *Spiel* 'game' is a superordinate level term. Firstly, if shown a particular type of game, e.g. board game or computer game, it is likely that a speaker would instinctively specify the type of game they see. It is relatively difficult to list numerous attributes which are common to all games (see 2.6), but significantly easier to do this for a specific genre of game. Similarly, listing individual parts of games in general is impossible without knowing the specific type of game, as is listing the ways in which someone might interact with one. The only feature which *Spiel* has in common with basic level terms, however, is the fact that children are likely to know the word *Spiel* and it is relatively short. For all other characteristics, however, *Spiel* seems to fit more readily on

the superordinate level. Similar reasoning can be provided for other nouns in the sample such as *Wissen* ‘knowledge’, *Kraut* ‘plant/herb’ and *Maß* ‘measure’ (see also Zubin & Köpcke 1986).

In total, 12 superordinates in the sample can be identified, meaning that it is now possible to determine whether OGAT can account for the genders of the nouns to which SUP.→*M,*F applies. The analysis shows that OGAT assigns the correct gender to the vast majority of these nouns, specifically 10/12 (83%), as exemplified by (13) and (14). The figure increases to 11/12 (92%) if the various adjustments to the theory proposed in 3.3 and 3.4 are incorporated into the model (see 3.5.2 for exceptions).

(13)	<i>Mittel</i> ‘means’	GENDER FEATURES SUP.→*M,*F		DEFAULT HIERARCHY		
				*N	*F	*M
	<i>der Mittel</i>	*!				*
	<i>die Mittel</i>	*!			*	
	☞ <i>das Mittel</i>			*		

(14)	<i>Fauna</i>	GENDER FEATURES SUP.→*M,*F -A→*M,*N		DEFAULT HIERARCHY		
				*N	*F	*M
	<i>der Fauna</i>	*	*!			*
	☞ <i>die Fauna</i>	*			*	
	<i>das Fauna</i>		*	*!		

3.2.2.3 DEVERBAL -T→*M,*N

The third constraint which is problematic in terms of its application is DEVERBAL -T→*M,*N “nouns formed by the suffixation of *-t*, e.g. *die Schrift*, cannot be masculine or neuter” (e.g. Flämig 1991: 453, Hickey 2000: 634, Hoberg 2004: 87, Chan 2005: 55). This is because the suffixation of *-t* is no longer productive in German (e.g. Hoberg 2004: 88, Fleischer & Barz 2012: 254). As a consequence, it is unclear to what extent the derivational pattern is transparent to native speakers of modern German, and therefore to which nouns the constraint DEVERBAL -T→*M,*N might realistically be applied.

Fleischer & Barz (2012: 254) argue that there are certain nouns, e.g. *Fahrt* ‘journey’ (from *fahren* ‘to travel’) and *Naht* ‘seam’ (from *nähen* ‘to sew’), whose verbal origin is clear from both a formal and semantic perspective, and others, e.g. *Schrift* ‘writing’ (from *schreiben* ‘to write’) and *Tracht* ‘constume/uniform’ (from *tragen* ‘to wear’), for which the association with the verb is less obvious due to the formal variation in the stem, but nevertheless recognisable. They identify a number of cases, however, in which – at least from a semantic perspective – the deverbal nature of the noun is relatively opaque, e.g. *Pflicht* ‘obligation’ (from *pflegen* ‘to care for/maintain’) and *Bucht* ‘bay’ (from

biegen 'to bend'), meaning that these are likely to be analysed synchronically as simplex (Fleischer & Barz 2012: 254).

In the sample, nouns to which DEVERBAL -T→*_M,*_N could potentially apply include *Schwestertracht*, *Durchschrift*, *Bildhauerkunst* and *Vertretungsvollmacht*. Both *Tracht* and *Schrift* are synchronically recognisable as deverbal derivatives according to Fleischer & Barz (2012: 254), as stated above, yet the deverbal origin of *Macht* 'power' (from *mögen* 'to like', see Kluge 2002: 453) and *Kunst* 'art' (from *können* 'to be able to', see Kluge 2002: 419) is opaque. It is therefore conceivable that the constraint DEVERBAL -T→*_M,*_N could apply to *Schwestertracht* and *Durchschrift*, but not to *Vertretungsvollmacht* or *Bildhauerkunst*.

Hoberg (2004: 88), however, argues that the suffixation of *-t* is not at all transparent in modern German and that all nouns which have been derived in this way are perceived as simplex. Hoberg (2004: 88) claims that the gender of these nouns is more likely to be accounted for synchronically through the association of feminine with the phonological ending /_(C) + f, x + t/, a principle also suggested by Köpcke (1982: 98), Zubin & Köpcke (1981: 440, 1984: 44), Mills (1986: 33), Köpcke & Zubin (1997), Menzel (2004: 68), Duden Grammatik (2009: 166) and Eisenberg (2013: 135). This exists as -FT/-CHT→*_M in the constraint set (see Chapter 2). Hoberg's (2004: 88) claim is strengthened by the observation that various other nouns which have the phonological ending /_(C) + f, x + t/ but are not of deverbal origin, e.g. *Frucht* 'fruit', have changed gender over time to become feminine (e.g. Hoberg 2004: 88, see also 2.8.2).

In terms of the sample, this phonological generalisation would account for the genders of *Tracht*, *Schrift* and *Macht*, but not *Kunst*. However, as stated above, the opacity of *Kunst* means that its gender cannot synchronically be accounted for by DEVERBAL -T→*_M,*_N either. A phonological association between final *-st* and feminine is also not possible due to the high proportion of masculine nouns in *-st* (approximately 65% based on nouns in the Duden Bibliothek-Express), e.g. *Ast*, *Test*, *Dienst*.

Therefore, since it is doubtful whether *-t* suffixation is always transparent to native speakers and also because the phonological generalisation is able to account for a higher proportion of nouns in the sample, the constraint -FT/-CHT→*_M is more favourable than DEVERBAL -T→*_M,*_N.

Of the nouns in the sample to which -FT/-CHT→*_M applies, OGAT in its current state is able to account for almost half. However, this figure increases to 80% if the various adjustments to the theory proposed in 3.3 and 3.4 are incorporated into the model (see 3.5.2 for exceptions).

3.2.2.4 -MA→*M,*F

The final problematic constraint is -MA→*M,*F “nouns (of Greek origin) with the “foreign” suffix *-ma*, e.g. *das Thema*, cannot be masculine or feminine” (e.g. Flämig 1991: 453, Weinrich 2007: 326, Duden Grammatik 2009: 166). This is problematic because *-ma* is not a synchronically productive suffix in German, appearing only on nouns borrowed from Greek. Crucially, it is not transparent, and is unlikely to be recognised as being present in nouns such as *Klima* and *Drama*, but as absent in nouns such as *Firma* (an Italian loan, see Duden Online-Wörterbuch 2018c).

The feature *-ma* can therefore only be analysed as a phonological ending. However, the principle “all nouns ending in *-ma* are neuter” does not meet Enger’s (2009) criteria. For instance, the principle fails the first criterion because, of the 98 nouns in *-ma* listed in the Duden Bibliothek-Express, almost half are not neuter, largely due to the strong association of final *-a* and feminine (e.g. Steinmetz 1986: 197, Steinmetz & Rice 1989: 166, Thomoglou 2004: 41, Duden Grammatik 2009: 164, Durrell 2011: 8, Kraiss 2014: 36, see also 2.3). This causes the principle to fail Yang’s (2005) Tolerance Principle.

Instead, in order to account for the neuter noun in the sample ending in *-ma*, *Sophisma*, the phonological constraint -ISMA→*M,*F might be hypothesised, which does fulfil Enger’s (2009) criteria. For instance, the constraint covers 83% of candidate nouns listed in the Duden Bibliothek-Express, thus fulfilling the first criterion (2.3), and also accounts for exceptions to the constraint -A→*M,*N, such as *Charisma*, *Schisma* etc., thereby also fulfilling the fifth criterion (2.7).

However, even with this constraint, OGAT in its current state is still unable to assign the correct gender to *Sophisma*, as shown in (15).

(15)

	GENDER FEATURES		DEFAULT HIERARCHY		
	-ISMA→*M,*F	-A→*M,*N	*N	*F	*M
<i>Sophisma</i> ‘sophism’					
<i>der Sophisma</i>	*	*!			*
☞ <i>die Sophisma</i>	*			*	
<i>das Sophisma</i>		*	*!		

In order for OGAT to produce the correct prediction, -ISMA→*M,*F must be ranked above -A→*M,*N. With this modification, the theory is then able to assign the correct gender to *Sophisma*, as shown in (16). This ranking would also account for other neuter nouns in the German lexicon in *-isma* such as *Prisma* and *Melisma* (as well as those mentioned above).

(16)

<i>Sophisma</i> 'sophism'	GENDER FEATURES		DEFAULT HIERARCHY		
	-ISMA → *M, *F	-A → *M, *N	*N	*F	*M
<i>der Sophisma</i>	*!	*			*
<i>die Sophisma</i>	*!			*	
☞ <i>das Sophisma</i>		*	*		

This ranking will be tested further in Chapter 4.

3.3 Noun-related issues

As stated in 3.1, in addition to constraint-related issues, there are also certain subsets of nouns in the sample which require closer consideration in order to establish whether OGAT can account for their genders in its current state, or whether any adaptations to the theory are needed. These are: loanwords (3.3.1), nouns with multiple meanings or forms (3.3.2), complex nouns (3.3.3) and the subset of 55 nouns (see 2.2), which includes proper nouns and MGNs (3.3.4). These will each be investigated in turn and a number of adjustments to OGAT will be proposed on the basis of the discussion.

3.3.1 Loanwords

There are a variety of loanwords in the sample which have been borrowed into German from a range of languages. Many of these are relatively old loans (e.g. from Latin or Greek), some of which have integrated into the German language to such an extent that they are no longer recognisable as loans to the majority of native German speakers, e.g. *Zins* (Latin *census*) and *Pfeffer* (Latin *piper*), the genders of which OGAT is able to correctly predict.

(17)

<i>Zins</i>	GENDER FEATURES			DEFAULT HIERARCHY		
	-C S MONO → *F, *N	CCVCC → *F, *N	-N(C)(C) MONO → *F, *N	*N	*F	*M
☞ <i>der Zins</i>						*
<i>die Zins</i>	*	*	*!		*	
<i>das Zins</i>	*	*	*!	*		

(18)

<i>Pfeffer</i>	GENDER FEATURES		DEFAULT HIERARCHY		
	-ER → *F, *N		*N	*F	*M
☞ <i>der Pfeffer</i>					*
<i>die Pfeffer</i>	*!			*	
<i>das Pfeffer</i>	*!		*		

Other, "learned" loans from Latin and Greek are more recognisable as loans to native speakers, but have integrated into the German gender system with such regularity that their gender can generally be predicted based on their form. As a result, there are many well-established GENDER FEATURES

constraints which relate specifically to the features of Latin and Greek loans, e.g. $-IUM \rightarrow *M, *F$, which applies to the loan in the sample *Antependium*, and $-ITIS \rightarrow *M, *N$, which applies to *Laryngitis*. As a result of these constraints, OGAT is able to correctly predict the genders of these nouns.

(19)

<i>Antependium</i>	GENDER FEATURES $-IUM \rightarrow *M, *F$	DEF. HIERARCHY		
		*N	*F	*M
<i>der Antependium</i>	*!			*
<i>die Antependium</i>	*!		*	
<i>☞ das Antependium</i>		*		

(20)

<i>Laryngitis</i>	GENDER FEATURES $-ITIS \rightarrow *M, *N$	DEF. HIERARCHY		
		*N	*F	*M
<i>der Laryngitis</i>	*!			*
<i>☞ die Laryngitis</i>			*	
<i>das Laryngitis</i>	*!	*		

Also in the sample are a number of more recent loans, many of which have been borrowed from English, e.g. *All-Star-Band*, *Bootcamp*. OGAT is able to account for the genders of many of these loans. However, there are a number of recent loans in the sample for which the theory in its current state is unable to account, as exemplified by *die Tramway* in (21). This suggests that recent loans may behave somewhat differently to older loans and native nouns with regard to gender assignment, indicating that some adaptations to OGAT are necessary.

(21)

<i>Tramway</i>	GENDER FEATURES	DEFAULT HIERARCHY		
		*N	*F	*M
<i>☞ der Tramway</i>				*
<i>die Tramway</i>			*!	
<i>das Tramway</i>		*!		

It has been the aim of various studies to determine how gender is assigned to recent loanwords in German, e.g. Arndt (1970), Carstensen (1980), Gregor (1983), Schlick (1984), Yang (1990), Kratochvílová (2000), Schulte-Beckhausen (2002), Chan (2005) and Onysko (2007). In these studies, a number of principles are claimed to account for the genders assigned to recent loans, which are thought to work alongside the meaning and form-based principles that are also applicable to native nouns (i.e. the current constraint set). These include: a principle assigning the gender of the loan in the donor language, a principle assigning the gender of the nearest semantic equivalent, a principle assigning the gender of a homonym/partial homonym, and the default gender principle.

In this section, each of these principles will be discussed in relation to the 45 relatively recent loanwords from the sample listed below, and it will be determined whether any of the additional

principles should be incorporated into OGAT, so that it is better able to account for the genders of recent loanwords.

<i>das Allegro</i>	<i>das Bootcamp</i>	<i>das Filet</i>	<i>der Rotang</i>
<i>die All-Star-Band</i>	<i>der Brandy</i>	<i>die Hairstylisten</i>	<i>die Rupiah</i>
<i>das Anderkonto</i>	<i>die Cantate/Kantate</i>	<i>das Indossament</i>	<i>die Steelband</i>
<i>der Ara</i>	<i>der Chickenwing</i>	<i>der Islam</i>	<i>das Stop-and-Go</i>
<i>das Arrival</i>	<i>das Coaching</i>	<i>der Key-Account-Manager</i>	<i>das Tamtam</i>
<i>der Audiostream</i>	<i>die Cranberry</i>	<i>der Koton</i>	<i>das Tempo</i>
<i>der/das Avis</i>	<i>der Defroster</i>	<i>der Mustang</i>	<i>der Test</i>
<i>der Aviso</i>	<i>der Dime</i>	<i>der/das New Look</i>	<i>das Tipi</i>
<i>das Aviso</i>	<i>der Docker</i>	<i>der New-Orleans-Jazz</i>	<i>die Tramway</i>
<i>das Bakschisch</i>	<i>der Dorfclub</i>	<i>der Presenter</i>	
<i>das Blackjack</i>	<i>das Eclair</i>	<i>der/das Renforcé</i>	
<i>die Bodybuilderin</i>	<i>der Encoder</i>	<i>die Rockoper</i>	

3.3.1.1 Donor gender

The first factor which is claimed to influence the gender of recent loanwords is the gender of the noun in the donor language (e.g. Gregor 1983: 23, Corbett 1991: 80, Schulte-Beckhausen 2002: 34-38). For instance, it might be claimed that the Italian loanword *Salto* is masculine in German due to the noun *salto* being masculine in Italian (Duden Grammatik 2009: 167f.).

This factor is, however, limited in its application, since it can of course only apply to nouns from donor languages which have a grammatical gender system which resembles that of German, e.g. Russian (M/F/N) and French (M/F), and cannot apply to loans from languages which do not, e.g. English and Turkish (no nominal gender). Because of this, incorporating this factor into OGAT will not help the theory account for any loans that it cannot currently account for which are from languages without nominal gender, e.g. English loans such as *Tramway* and *Bootcamp*.

Additionally, Gregor (1983: 23) and Schulte-Beckhausen (2002: 34f.) highlight that this factor is also limited because it requires speakers to have considerable knowledge of the donor language, and is therefore only available to a small proportion of speakers. Schulte-Beckhausen (2002: 35f.) observes that the borrowing of gender along with a noun tends to be confined to the very early stages of borrowing, and that as the loan becomes increasingly used among monolingual German speakers, the gender of the noun in the donor language ceases to be influential. She claims that factors such as formal characteristics or semantic equivalency (see 3.3.1.2) are more likely to determine the gender of a loan, as these are available to all speakers at all stages of the borrowing process.

In order to assess whether gender borrowing is likely to influence the gender of the recent loanwords in the sample, the table below has been compiled, showing the loans from donor languages with

nominal gender, their donor language (as recorded in the Duden) and donor gender, alongside an indication of whether the outcome of OGAT would be positively affected by the inclusion of this principle.

Loanword	Donor gender	Same gender?	OGAT outcome positively affected?
<i>die Cantate/Kantate</i>	F (Italian)	✓	x
<i>die Rockoper</i>	F (Italian)	✓	x
<i>der Ara</i>	M (French)	✓	x
<i>der Aviso</i>	M (French)	✓	x
<i>der Koton</i>	M (French)	✓	x
<i>der Islam</i>	M (Arabic)	✓	x
<i>der/das Avis</i>	M (French)	✓/x	x
<i>das Eclair</i>	M (French)	x	x
<i>das Filet</i>	M (French)	x	x
<i>das Indossament</i>	M (French)	x	x
<i>das Tamtam</i>	M (French)	x	x
<i>das Anderkonto</i>	M (Italian)	x	x
<i>das Aviso</i>	M (Italian)	x	x
<i>das Tempo</i>	M (Italian)	x	x

The table largely confirms Gregor (1983) and Schulte-Beckhausen (2002)'s claim that the gender of the loan in the donor language is unlikely to be a synchronically influential factor in determining the gender of the loan in the recipient language. Not only are there only six cases in the sample in which the German gender of the loan is the same as the donor gender, but in no case would incorporation of the principle positively affect the OGAT outcome. That is to say, the donor gender principle does not help OGAT yield the correct prediction in cases where it currently cannot.

This, coupled with the fact that it is unlikely to form part of a native German speaker's gender assignment system, since it requires considerable knowledge of the donor language which most native German speakers will not have, means that the influence of gender borrowing will not be incorporated into OGAT.

3.3.1.2 Semantic equivalence

Another factor which is argued to have an impact on the gender of loanwords in German is the gender of the nearest semantic equivalent in the recipient language, e.g. *Airport* might be said to be assigned masculine due to the semantic equivalent *der Flughafen* (e.g. Clyne 1967: 42, Arndt 1970: 248, Geerts 1975: 118f., Carstensen 1980: 15ff., Gregor 1983: 59, Schlick 1984: 404, Yang 1990: 153f., Hickey 2000: 639f., Kratochvílová 2000: 68ff., Schulte-Beckhausen 2002: 65-68, Chan 2005: 103-110, Onysko 2007). In the sample, this could potentially explain the gender of loans for which OGAT in its current state cannot account, such as *das Bootcamp* (*das Lager*) and *die Tramway* (*die Straßenbahn*). Both Gregor (1983) and Schulte-Beckhausen (2002: 65-68) state that this factor might be especially important for

loans which have no other features that are relevant to gender. In the sample, this appears to be true for loans such as *das Bakschisch* (*das Trinkgeld*) and *das Tipi* (*das Indianerzelt*).

A number of restrictions on semantic equivalence have, however, been identified. The first is that sometimes there is no German equivalent to a loan, since borrowings often represent novel concepts (Onysko 2007: 166). This can be seen in the sample with e.g. *Jazz*.⁶ This alone is not necessarily problematic for OGAT, since an optimal candidate can be found regardless of whether or not there are any relevant GENDER FEATURES. Moreover, many of the loans for which the theory cannot currently account do have a semantic equivalent.

Secondly, there are a number of loans which do not appear to adopt the gender of the German equivalent, even when the equivalent is direct, e.g. *das Ticket* (*die Fahrkarte*) and *der Speed* (*die Geschwindigkeit*) (e.g. Carstensen 1980, Onysko 2007: 166f.). In the sample, this seems to be true for loans such as *das Konto* (*die Bankverbindung*). Again, this is also not necessarily problematic for OGAT, as the gender of these loans might be explained as a consequence of the interactions of other GENDER FEATURES constraints.

The third criticism of semantic equivalency is that there are often multiple possible equivalents to a loanword and it is not always clear which equivalent native speakers perceive to be the closest (Carstensen 1980: 16, Schulte-Beckhausen 2002: 67, Chan 2005: 104f., Onysko 2007: 166f.). For instance, in the sample, *der/das Avis* has a number of equally plausible equivalents, such as *der Hinweis*, *die Anzeige*, *die Benachrichtigung*, *die Mitteilung* etc. Similarly, *das Tempo* could be seen as equivalent to *die Geschwindigkeit*, *die Schnelle*, *der Rhythmus* or *der Takt*. One possible solution to this problem is employed by Hunt (2016), who in his study calculates which is the most frequent semantic equivalent in cases where there are multiple options. Additionally, it has been observed that the gender of a semantic equivalent is likely to be more influential if the equivalent is also a transparent cognate of the loanword, e.g. *der Background* – *der Hintergrund*, *der Market* – *der Markt* (Geerts 1996, Fuller & Lehnert 2000: 413, Chan 2005: 75, Onysko 2007: 168).

An advantage of incorporating semantic equivalency into a model of gender assignment has been highlighted by Schulte-Beckhausen (2002: 67, 116), who states that it has the potential to account for multiple-gender loans. For instance, the multiple-gender status of *die/der Browning* could be explained by the competition between the two equivalents *die Pistole* and *der Revolver*, and *der/das Karibu* might be explained by the competing equivalents *der Hirsch* and *das Ren* (Schulte-Beckhausen 2000: 116). In the sample, semantic equivalency could provide an explanation for the variable gender

⁶ Note that *Musik/Musikrichtung* is a hypernym rather than a semantic equivalent – see 1.3.3.3.

of *der/das New Look*, due to the differing genders of the two possible equivalents *der (neue) Stil* and *das (neue) Aussehen*. For further discussion of MGNs, see 3.3.4.8.

Importantly, Schulte-Beckhausen (2002: 67f.) observes that where a loanword is a compound, the semantic equivalence principle seems to often apply to the compound as a whole, rather than to just the right-hand element as might be expected (see 3.3.3). For example, Schulte-Beckhausen (2002: 67) claims that *Arrowroot* is neuter because of the equivalent of the whole compound, *das Stärkemehl*, and not feminine because of the equivalent of the right-hand element, *die Wurzel*. This also seems true for certain loans in the sample, e.g. *Tramway*, which is feminine seemingly due to the equivalent *die Straßenbahn*, rather than masculine due to *der Weg*, or, for that matter, any other features of *Way* (e.g. CV→*N; V_{LONG}(C)→*F).

This suggests that some compound loanwords are not analysed into their original constituent parts by German speakers, which makes sense, given that to do so requires knowledge of the donor language, which not all speakers have. Indeed, neither *Root* nor *Way* are listed in the Duden, indicating that they are not established lexemes in German, and it is thus unlikely that native speakers will analyse compounds of which they are the right-hand element into their constituent parts. Consequently, the gender of such compounds will be based on the features of the compound as a whole rather than the features of just the right-hand element.

Therefore, compound loans whose right-hand element is not listed in the Duden and is thus unlikely to be recognised by most native German-speakers should be analysed as wholes in OGAT. The table below shows the 17 compound loanwords in the sample, along with information as to whether their right-hand element is in the Duden.

Compound	RH element in Duden?
<i>die All-Star-Band</i>	✓
<i>die Steelband</i>	✓
<i>das Anderkonto</i>	✓
<i>das Bootcamp</i>	✓
<i>die Hairstylistin</i>	✓
<i>der Key-Account-Manager</i>	✓
<i>der/das New Look</i>	✓
<i>der New-Orleans-Jazz</i>	✓
<i>die Rockoper</i>	✓
<i>der Dorfclub</i>	✓

Compound	RH element in Duden?
<i>der Audiostream</i>	x
<i>das Blackjack</i>	x
<i>die Bodybuilderin</i>	x
<i>der Chickenwing</i>	x
<i>die Cranberry</i> ⁷	x
<i>das Stop-and-Go</i>	x
<i>die Tramway</i>	x ⁸

⁷ *Cranberry* is technically not a compound since *cran-* is a “cranberry”-morpheme. Nevertheless, it has been included here because *berry* is a stand-alone unit in English which could plausibly exist as a loan in German.

⁸ The Duden provides further evidence that *way* does not serve as a gender-determining element, since, of the 11 other compounds in *-way*, six are masculine (e.g. *Broadway*, *Highway*), two are feminine (*Gangway*, *Subway*), two are neuter (*Fairway*, *Gateway*) and one is a MGN (*der/die Runway*). Yang (1990: 154) and Onysko (2007: 177) also maintain that *-way* compounds are analysed as simplex loans.

The results indicate that 10 of the compounds are likely to be analysed into their original constituent parts and 7 are not. This means that the principle of semantic equivalency would apply to the right-hand element only in these 10 cases and to the whole compound in the remaining 7 cases.

In order to assess the relevance and accuracy of the semantic equivalence principle, the table below has been compiled, showing the 45 relatively recent loanwords from the sample along with their nearest semantic equivalent(s). For compound loans, the gender-determining element is underlined (e.g. *Tramway* vs. *Steelband*). For loans with multiple plausible equivalents, the most frequent equivalent according to the Leipzig Corpora Collection is listed in bold, and equivalents which are cognates are listed in italics.

	Loanword	Semantic equivalent(s)	Same gender?
No true equivalent	das Allegro	–	–
	der Ara	–	–
	der Aviso	–	–
	die Cantate/Kantate	–	–
	der Islam	–	–
	der Mustang	–	–
	der <u>New-Orleans-Jazz</u>	–	–
	der/das Renforcé	–	–
	die <u>Rockoper</u>	–	–
	der Rotang	–	–
	die Rupiah	–	–
das <u>Stop-and-Go</u>	–	–	
One equivalent	das <u>Anderkonto</u>	die Bankverbindung	x
	das Bakschisch	das Trinkgeld	✓
	das <u>Blackjack</u>	das Siebzehndvier	✓
	die <u>Bodybuilderin</u>	die Muskelfrau	✓
	das <u>Bootcamp</u>	das Lager	✓
	der <u>Chickenwing</u>	der Hähnchenflügel	✓
	der Dime	der Cent	✓
	das Tipi	das Indianerzelt	✓
Multiple possible equivalents of the same gender	der Brandy	<i>der Branntwein</i> , der Weinbrand	✓
	die <u>Cranberry</u>	die Preiselbeere , <i>die Kronsbeere</i>	✓
	der Docker	der Hafenarbeiter , der Schauermann	✓
	die <u>Hairstylistin</u>	<i>die Stilistin</i> , die Künstlerin	✓
	der <u>Key-Account-Manager</u>	der Geschäftsführer, der Leiter, der Chef	✓
	der Presenter	der Moderator, der Showmaster, der Sprecher , der Ansager, <i>der Präsentator</i>	✓
Multiple possible equivalents of different genders	die <u>Tramway</u>	die Straßenbahn , die Elektrische	✓
	die <u>All-Star-Band</u>	die Gruppe , die Kapelle, das Ensemble	✓
	die <u>Steelband</u>	die Gruppe , die Kapelle, das Ensemble	✓
	der <u>Audiostream</u>	der Datenstrom , <i>der Audiostrom</i> , die Audiodatei, der/das Audiofile	✓
	der Defroster	der Enteiser , <i>der Entfroster</i> , die Enteisungsanlage	✓
der Dorf <u>club</u>	der Verein , die Vereinigung, die Gemeinschaft, der Verband, der Bund	✓	

der Encoder	der Geber , <i>der Codierer</i> , der Verschlüsseler, das Codiergerät, die Verschlüsselungseinrichtung	✓
das Filet	das Steak , das Lendenstück, das Kotelett, die Lendenschnitte	✓
der/das Avis	der Hinweis , <i>das Aviso</i> , die Anzeige, die Benachrichtigung, die Ankündigung, die Mitteilung, die Meldung	✓
der/das New <u>Look</u>	der Stil , das Aussehen	✓
der Koton	<i>der/das Cotton</i> , die Baumwolle	✓*
das Tamtam	der/das Gong , das Becken	✓*
das Aviso	<i>der/das Avis</i> , der Hinweis , die Anzeige, die Benachrichtigung, die Ankündigung, die Mitteilung, die Meldung	✓*
das Arrival	die Ankunft , die Anreise, das Eintreffen, das Erscheinen	x
das Coaching	das Trainieren, das Betreuen, das Beraten, die Betreuung , die Beratung	x
das Eclair	der Liebesknochen , die Hasenpfote, die Kaffeestange	x
das Indossament	der Vermerk, die Übertragung	x
das Tempo	die Geschwindigkeit , die Schnelle, Schnelligkeit, der Rhythmus, Takt	x
der Test	die Prüfung, Klausur, die Probe, die Untersuchung , das Examen, der Versuch	x

The table shows that 12 loanwords in the sample have no true equivalent in German and seven loanwords do not share their gender with their nearest, most frequent equivalent. The gender of these nouns must therefore be ultimately determined by other constraints. For 26 loans in the sample, however, the gender of the nearest, most frequent equivalent, or the equivalent which is also a cognate, was found to be the same as their own, suggesting that semantic equivalence could be a gender-determining factor for these nouns. In three cases, the most frequent or cognate equivalent is a MGN and shares one of its possible genders with the loan.

Importantly, in terms of OGAT, incorporation of the semantic equivalence principle would enable the theory to account for the genders of loans in the sample for which it currently cannot account, e.g. *Bootcamp* and *Bakschisch*, as shown in (22) and (23).

(22a)

<i>Bootcamp</i>	GENDER FEATURES			DEFAULT HIERARCHY		
	-N(c)(c) MONO→*F,*N	DWELL→*M	CVCC→*F	*N	*F	*M
☞ <i>der Bootcamp</i>		*				*
<i>die Bootcamp</i>	*		*!		*	
<i>das Bootcamp</i>	*			*!		

(22b)

	GENDER FEATURES				DEFAULT HIERARCHY		
	SEM. EQUIV. → *M, *F (LAGER)	-N(C)(C) MONO →*F, *N	DWELL →*M	CVCC →*F	*N	*F	*M
<i>Bootcamp</i>							
<i>der Bootcamp</i>	*		*!				*
<i>die Bootcamp</i>	*	*		*!		*	
<i>das Bootcamp</i>		*			*		

(23a)

	GENDER FEATURES		DEFAULT HIERARCHY		
			*N	*F	*M
<i>Bakschisch</i>					
<i>der Bakschisch</i>					*
<i>die Bakschisch</i>				*!	
<i>das Bakschisch</i>			*!		

(23b)

	GENDER FEATURES		DEFAULT HIERARCHY		
	SEM. EQUIV. →*M, *F (TRINKGELD)		*N	*F	*M
<i>Bakschisch</i>					
<i>der Bakschisch</i>		*!			*
<i>die Bakschisch</i>		*!		*	
<i>das Bakschisch</i>			*		

If semantic equivalency is incorporated as an equally-ranked GENDER FEATURES constraint, for most loans this either affects the outcome positively, as in (22) and (23), or does not affect it at all. However, for a small number of nouns, e.g. *Arrival*, this causes OGAT to yield the incorrect prediction where it originally did not.

(24)

	GENDER FEATURES		DEFAULT HIERARCHY		
	SEM. EQUIV. → *M, *N (ANKUNFT)	-AL ENG. →*M, *F ⁹	*N	*F	*M
<i>Arrival</i>					
<i>der Arrival</i>	*	*!			*
<i>die Arrival</i>		*		*	
<i>das Arrival</i>	*		*!		

This problem can be simply solved through ranking. If the semantic equivalency constraint is ranked below all other GENDER FEATURES constraints, the theory can still account for the loans for which it originally could not account, e.g. *Bootcamp* and *Bakschisch*, whilst avoiding any negative impact on loans which were originally unproblematic, e.g. *Arrival*.

(25)

	GENDER FEATURES		DEFAULT HIERARCHY		
	-AL ENG. →*M, *F	SEM. EQUIV. → *M, *N (ANKUNFT)	*N	*F	*M
<i>Arrival</i>					
<i>der Arrival</i>	*!	*			*
<i>die Arrival</i>	*!			*	
<i>das Arrival</i>		*	*		

⁹ See 3.4.1 for details of this constraint.

(26)

<i>Bootcamp</i>	GENDER FEATURES			DEFAULT HIERARCHY			
	-N(C)(C) →*F,*N	DWELL →*M	CVCC →*F	SEM. EQUIV.→ *M,*F (LAGER)	*N	*F	*M
<i>der Bootcamp</i>	*			*!			*
<i>die Bootcamp</i>	*		*!	*		*	
<i>☞ das Bootcamp</i>	*				*		

(27)

<i>Bakschisch</i>	GENDER FEATURES		DEFAULT HIERARCHY		
		SEM. EQUIV.→*M,*F (TRINKGELD)	*N	*F	*M
<i>der Bakschisch</i>		*!			*
<i>die Bakschisch</i>		*!		*	
<i>☞ das Bakschisch</i>			*		

In sum, it has been demonstrated that the data justify the insertion of a semantic equivalence constraint which is ranked lower than other GENDER FEATURES constraints in OGAT.

3.3.1.3 Homonymy/partial homonymy

The third possible additional factor which may influence the gender of loanwords is the gender of a corresponding German homonym or partial homonym. It is argued that this could consist of a phonologically and/or orthographically similar noun, which does not necessarily have to be semantically equivalent or etymologically related to the loan (Clyne 1967: 42; Arndt 1970: 249, 252; Gregor 1983: 77; Schlick 1984: 409-417; Schulte-Beckhausen 2002: 59f., 63ff.).

Schulte-Beckhausen (2002: 59), for instance, claims that the gender of several loanwords is influenced by the gender of a phonologically similar German noun, e.g. *die Party* (celebration) ← *die Partie* 'section'. It is also argued that the gender of a loan can be influenced by multiple phonologically similar nouns, e.g. *der Snack* ← *der Dreck/Fleck/Scheck/Schreck* (Schulte-Beckhausen 2002: 60). Sometimes the phonologically similar nouns which are claimed to influence the gender of a loan are also semantically and/or etymologically related to the loan, e.g. *die Army* ← *die Armee* (Schlick 1984: 411, 417; Schulte-Beckhausen 2002: 60).

Graphemically similar nouns (that may or may not also be phonologically similar) are also claimed to influence the gender of loans. Schulte-Beckhausen (2002: 63ff.) states that this is most likely to occur when the loan is used more in written communication than speech. Examples of this might include: *der Mull* [mʊl] (fabric) ← *der Müll* [mʏl] 'rubbish' and *der Trouble* [ˈtrʌb] ← *der Trubel* [ˈtru:b] 'hubbub' (Schlick 1984: 411, Schulte-Beckhausen 2002: 63).

Onysko (2007: 153), however, disputes the claim that homonymic relations influence the gender of loans. He argues that cases which are said to demonstrate the homonymy principle can usually be explained by either semantic equivalence or form/meaning-based principles. This indeed seems to be the case for a number of examples cited above, such as *der Mull*, which can instead be explained by the semantic constraint FABRIC → *F, *N. Consequently, Onysko (2007: 153) maintains that any instances of homonyms sharing a gender arise only by coincidence. He lists numerous examples of loans being assigned a different gender to that of their corresponding homonym, e.g. *der Comic* 'comic strip' – *die Komik* 'humour', *der Doc* 'doctor' – *das Dock* 'dock', *die Page* 'page' – *der Page* 'footballer', *die Gang* 'group' – *der Gang* 'corridor/alley', *die Wall* 'wall' – *der Wall* 'rampart'.

As for the loans in the sample, the homonymy principle could perhaps be said to apply to nouns such as *der Dime* ← *der Darm/Damm/Dom* and *das Arrival* ← *das Revival*. However, in both of these cases, the gender of the loan can more plausibly be explained using already existing constraints, namely MON. UNITS → *F, *N and -AL ENG. → *M, *F respectively. Moreover, there are many cases in the sample in which the loan does not assume the gender of its corresponding (partial) homonym(s), e.g. *die All-Star-Band* – *der Band/das Band*; *der Ara* – *die Ära*; *die Rockoper* – *der Ober/das Opfer*. Consequently, inclusion of the homonymy principle would generally not positively affect OGAT's ability to account for the sample data.

Not only does the homonymy principle seem ineffective, but the vagueness of partial homonymy is also problematic. That is to say, it is unclear to what extent the partial homonym must resemble the loanword in order for its gender to be an influential factor. Would, for instance, *der Jasmin* be classified as a partial homonym of *der Jazz*, or *der Salm* as a partial homonym of *der Islam*? Furthermore, where the gender of a partial homonym might be said to influence the gender of a loan, there often exist other partial homonyms of different genders. For instance, the gender of *der Audiostream* might be said to be influenced by the gender of the partial homonym *der Strom*, but, according to the homonymy principle, it could also be influenced by the gender of another partial homonym *die Strieme*. It is instead more plausible that *der Strom* would influence the gender of *der Audiostream* due to semantic equivalence (3.3.1.2). Similarly, *der Test* might be argued to have masculine gender because of partial homonyms such as *der Rest*. Again, if this is the case, the gender of *Test* could equally be influenced by the gender of the partial homonyms *das Fest*, *das Nest* and *die Pest*.

For the above reasons, homonymy/partial homonymy will not be incorporated into OGAT as a factor which could determine the gender of a loanword.

3.3.1.4 Default gender

A final possibility for the assignment of gender to loanwords is that loans automatically receive a default gender. Poplack, Pousada & Sankoff (1982: 21ff.) state that it is possible that loans are simply assigned the most unmarked gender in the recipient language, assuming this to be the gender to which the largest number of nouns belong (for German this is generally agreed to be masculine, see 1.4.5.3). To assess the validity of this claim, the distribution of genders among various samples of loanwords, including loans from the present study, is shown in Table 4.

Source		Sample size	M	F	N	Variable gender
Arndt (1970)	Nouns frequently used in US English tested on 25 native German speakers	130	35.0%	50.0%	15.0%	-
Yang (1990)	English loanwords in 24 issues of <i>Der Spiegel</i> (1950-1980)	1204	60.1%	16.3%	23.6%	-
OWID (1990-2017)	Neologism database including many loans (1990-2017)	1466	37.7%	28.1%	34.2%	8.1% (not included in total calculations)
Schulte-Beckhausen (2002)	English monosyllables found in various German dictionaries (1852-1999)	402	53.5%	7.7%	18.7%	20.1% (14.2% are M/N)
Chan (2005)	English loans in <i>Duden: Das große Wörterbuch der deutschen Sprache</i> (1999)	3105	48.4%	16.0%	29.5%	6.2% (4.5% are M/N)
Onysko (2007)	English loans in <i>Der Spiegel corpus</i> (2000)	1023	61.8%	17.9%	19.7%	0.6%
The present study	Recent loans from a systematically-selected sample from the <i>Duden Rechtschreibung</i> (2013)	45	37.8%	22.2%	33.3%	6.7% (all are M/N)

Table 4: The distribution of genders among loanwords according to various studies

It is clear from Table 4 that not all loans are assigned masculine, or any one gender in German, thereby eliminating the possibility of an automatic default. However, it is interesting to observe that masculine is the largest category in all but one study. Carstensen (1980) and Callies et al. (2012: 68) also report finding the majority of loanwords to be masculine in their studies, despite not giving exact figures. This data is consistent with Rice's (2006: 1406) argument that masculine is at the top of OGAT's default markedness hierarchy in German (*N»*F»*M) as it is the category to which most nouns belong, i.e. is the "stochastic default" (Köpcke & Zubin 2009: 148ff.) (see 1.4.5.3).

However, if the justification for the default hierarchy is indeed category size, then the evidence from loans in Table 4 does not support a *N»*F»*M hierarchy, but rather the hierarchy *F»*N»*M, since in almost all cases there are more neuter than feminine loans. As discussed in 1.4.5.3, this ranking is

also supported by data from e.g. Augst (1975) and Köpcke (1982). Testing of alternative default hierarchies will take place in 3.4.3.

3.3.1.5 Summary

In summary, it has been demonstrated that OGAT in its original state is able to account for many loanwords in the sample, including nouns which are synchronically unlikely to be perceived as loans, “learned” loans, as well as most recognisably recent loans. However, there are a number of recognisably recent loans in the sample for which OGAT in its original state cannot account. Consequently, various additional factors claimed to account for the genders of loanwords have been explored in order to determine whether their incorporation into OGAT is justifiable.

It was found that donor gender, homonymy/partial homonymy and default gender are unlikely to be part of OGAT due to their inability to account for the data, as well as more specific problems such as the unavailability of the donor gender principle to most native speakers and the vagueness of the homonymy/partial homonymy principle.

The semantic equivalence principle, however, was shown to be effective in accounting for the sample data. It was demonstrated that the principle is most effective for all loans in the sample if the semantic equivalence constraint is ranked lower than other GENDER FEATURES. With this adaptation, OGAT can account for the genders of a higher proportion of loans in the sample than the original version of the theory, including loans such as *Tramway* and *Bootcamp*, for which the original version is unable to account.

3.3.2 Nouns with multiple meanings or forms

Another subset of nouns in the sample which require closer consideration in order to determine whether or not OGAT is able to account for their genders are nouns with multiple possible meanings or forms. For a number of nouns in the sample of this type, their different meanings or forms require the application of different constraints, meaning that in some cases OGAT would predict different genders for the different meanings or forms of the same noun. This then raises the question of whether we do in fact consider the different meanings and forms to be variants of the same lexical entry or instead as relating to different nouns. Nouns with multiple meanings will be considered in 3.3.2.1 before discussing nouns with variable forms in 3.3.2.2.

3.3.2.1 Meaning variation

One set of nouns of which there are multiple variants which could prove problematic for OGAT are the nouns in the sample with multiple possible meanings or “senses”. These include nouns such as *Läufer*, which can refer to a runner [sport], a rug, a bishop [chess] or a rotor, and *Tamtam*, which can refer to a musical instrument or a general commotion. Section 3.3.2.1.1 will discuss how such nouns may be differentiated from nouns with a single sense, and these methods will then be applied to nouns in the sample in 3.3.2.1.2. Section 3.3.2.1.3 will subsequently consider the distinction between cases of polysemy and homonymy, and this will be discussed in terms of the sample and OGAT in 3.3.2.1.4. Finally, various approaches to polysemy will be examined in Section 3.3.2.1.5, before considering these in relation to the sample and OGAT in 3.3.2.1.6.

3.3.2.1.1 Multiple meanings vs. one meaning

It is claimed that nouns with multiple senses can be distinguished from monosemous nouns via a series of tests (e.g. Cruse 1986, Geeraerts 1993). The use of nouns with multiple senses is said to lead to ambiguity, whereas the use of monosemous nouns is said to lead only to vagueness (e.g. Steen 2007: 150, Murphy 2010: 84f.). For instance, *I went to the bank* is ambiguous, as it is unclear whether speaker is referring to the financial establishment or the riverside, and the hearer must select one of the senses in order for successful communication to occur (e.g. Cruse 1986: 51, Steen 2007: 150). The phrase *I saw my cousin*, by contrast, is vague, because it is unclear whether the speaker saw a female or male, maternal or paternal cousin, but the hearer does not need to determine this in order to comprehend the utterance (e.g. Cruse 1986: 51, Steen 2007: 150). Nouns with multiple senses and monosemous nouns can therefore be distinguished by testing for ambiguity. This can be done in a number of ways, including: the definition test, contrast tests and the zeugma test (e.g. Steen 2007: 151f., Murphy 2010: 85ff.).

The definition test assesses whether a term can be easily defined using a single definition or not (e.g. Cruse 1986: 51, Murphy 2010: 85). The term *cousin*, for instance, can be defined relatively easily as ‘child of an uncle/aunt’ – a definition which covers all possible referents of the term (male/female, maternal/paternal etc.). For *bank*, however, no single definition can be stated which covers both the interpretations ‘riverside’ and ‘financial establishment’. According to the definition test, *cousin* is therefore vague, and *bank* is ambiguous.

The contrast test involves determining whether different senses of a word can co-occur in a construction without producing infelicity. If they can, the word is said to have contrasting senses and thus be ambiguous (Geeraerts 1993, Steen 2007: 151f., Murphy 2010: 85). One example of this would

be testing whether two senses of a word can be used in a conjoined positive and negative assertion without the assertion being contradictory, as exemplified by (28a) and (28b) (Quine 1960: 129, Steen 2007: 151).

(28a) He's going to the bank_[RIVER] but he's not going to the bank_[INSTITUTION]

(28b) #We met her cousin_[MATERNAL] but we didn't meet her cousin_[PATERNAL]

(28a) uses different senses of *bank*, yet is not contradictory. These senses of *bank* are therefore said to contrast, demonstrating that *bank* is ambiguous. In (28b), however, the use of the different senses of *cousin* leads to a contradiction, suggesting that *cousin* is not ambiguous.

A final way of distinguishing between vagueness and ambiguity is the “zeugma” test (e.g. Cruse 1986, 2000; Pustejovsky 1995; Copestake & Briscoe 1996). This consists of constructing phrases to try and create zeugmas, i.e. utterances which are semantically odd due to distinct senses of a word being activated simultaneously. If a zeugmatic effect is produced, it is said that the word is ambiguous (e.g. Cruse 1986: 63, Murphy 2010: 86f., Urgelles-Coll 2010: 109). For instance, (29a) gives rise to zeugma due to the incompatible coordination of the sense of *case* ‘box’ and the sense ‘instance of disease’, indicating that *case* is ambiguous.

(29a) #She had a case of wine and another of pneumonia.

By contrast, (29b) does not produce a zeugmatic effect, despite *cousin* referring to a male cousin and *another* referring to a female cousin.

(29b) She has a cousin who is a businessman and another who is an actress.

This suggests that *cousin* has just one sense which is not ambiguous but vague.

The above examples demonstrate that ambiguity tests can be extremely useful when identifying monosemy. However, it is important to note that these tests do not always produce conclusive results (e.g. Cruse 1986, Geeraerts 1993, Lewandowska-Tomaszczyk 2007: 141-144). It has, for instance, been observed that the results of the zeugma test are sometimes context dependent (Norricks 1981: 115, Cruse 1986: 65f.), and that the definition test can be unreliable given that it relies on an individual's ability to produce definitions, which are notoriously difficult to construct (e.g. Laurence & Margolis 1999: 14). Consequently, it is possible that, for some items, the tests will not always produce the same results (e.g. Geeraerts 1993, Lewandowska-Tomaszczyk 2007: 143f.).

3.3.2.1.2 Multiple meanings vs. one meaning in the sample

In terms of the sample, the ambiguity tests described above can be successfully used to identify the nouns with multiple senses. For example, the results below clearly demonstrate the ambiguous nature of *Läufer*, indicating that it has multiple senses, and the vagueness of *Enkelin*, indicating that it has just one sense.

Definition test:	(30a) Enkelin = 'Tochter des Kindes [daughter of one's child]' (30b) Läufer = 'jemand, der das Laufen als Sport betreibt [somebody who runs as a sport]' and 'Schachfigur [chess piece]'
Contrast test:	(31a) #Meine Enkelin _[KIND DES SOHNES] sah ich nicht, aber meine Enkelin _[KIND DER TOCHTER] schon. '#I didn't see my granddaughter _[SON'S CHILD] , but I did see my granddaughter _[DAUGHTER'S CHILD] .' (31b) Den Läufer _[SPORT] sah ich nicht, aber den Läufer _[SCHACH] schon. 'I didn't see the runner _[SPORT] , but I did see the bishop _[CHESS] '.
Zeugma test:	(32a) Ich habe eine Enkelin _[KIND DES SOHNES] , die 21 ist und eine _[KIND DER TOCHTER] , die 5 ist. 'I have one granddaughter _[SON'S CHILD] who is 21 and one _[DAUGHTER'S CHILD] who is 5.' (32b) #Es gibt einen Läufer _[SPORT] im Olympiastadion und einen _[SCHACH] in der Spielzeugkiste. '#There is a runner _[SPORT] in the Olympic stadium and one _[CHESS] in the toy box'.

According to the tests, there are also quite a number of other nouns in the sample whose use leads to ambiguity, indicating that they have multiple senses. These include: *Knacker*, *Tamtam*, *Marke*, *Perle*, *Klasse*, *Leiter*, *Band*, *Hirsch*, *Cranberry*, *Kartoffel*, *Juchten*, *Ruder*, *Rotang* and *Alphabet*. There do not seem to be any obvious borderline cases in the sample for which the three tests produce differing results.

3.3.2.1.3 Polysemy vs. homonymy

Now that a distinction has been made between words whose use leads to vagueness and those whose use leads to ambiguity, it is now necessary to establish a distinction between two different sources of ambiguity: polysemy and homonymy.

It is generally agreed that homonymy occurs when a word form has two distinct senses which are not clearly related to each other, such as *bank*_[RIVER] and *bank*_[INSTITUTION] (e.g. Cruse 1986: 80, Blank 2003: 273, Steen 2007: 141). Homonyms are said to be linked to independent lexemes which, synchronically, have the same surface form by coincidence. Polysemy, by contrast, is said to occur when a single lexeme has multiple distinct senses that are related to one another, which arises through some form of extension of the original sense (e.g. Cruse 1986: 49, Löbner 2002: 60, Taylor 2003: 144). These distinct senses become conventionalised, and are stable across contexts, as opposed to indirect

meanings which arise from the ad hoc use of a word to refer to something in a particular context (e.g. Cruse 1986: 50, Steen 2007: 139f., Murphy 2010: 89).

There are three main ways suggested in the literature to distinguish between homonymy and polysemy. The first method is to examine the etymology of the items. Homonyms are said to typically have divergent etymologies, whereas polysemes typically derive from same source (e.g. Croft & Cruse 2004: 111, Lewandowska-Tomaszczyk 2007: 142).

While the etymological criterion can be extremely useful, it is not always the case that the etymology aligns with present-day speakers' perception of the relation between two items, which is the concern of the present study (e.g. Blank 2003: 276, Croft & Cruse 2004: 111). There are, for instance, multiple examples of originally polysemic forms whose senses have been reanalysed as being semantically unrelated and are therefore synchronically perceived as homonyms. Blank (2003: 276) provides the examples of *Schloss*₁ 'lock' vs. *Schloss*₂ 'castle', shown in (33), which are usually perceived as homonyms, yet both derive from the same source.

(33) MHG *sloz* 'lock' [>METAPHOR> 'castle locking a valley or a pass' >EXTENSION>] 'castle, palace'
(hence ModG *Schloss*₁ 'lock' *Schloss*₂ 'castle')

Additionally, there are numerous cases of originally homonymic forms which are analysed synchronically as being semantically related and therefore as polysemous. Langacker (1987: 387), for example, states that many speakers nowadays consider the meaning of *ear of corn* as a semantic extension of *ear* as a body part, despite these having divergent etymologies. Etymology is therefore unreliable as it is overwritten by reanalysis.

A second way to distinguish between homonymy and polysemy is to examine possible semantic relationships between two senses. If there is no conceivable relation between them, the items are homonymous, but if one sense can be considered an extension of the other, they are polysemous (e.g. Blank 2003: 275, Croft & Cruse 2004: 111, Murphy 2010: 91). Specifically, Blank (2003: 270-5) argues that polysemy can be identified if the conventionalised senses in question are linked by one of seven synchronic relations, including: metaphor (e.g. *snake* 'animal', 'deceitful person') and autohyponymy (e.g. *drink* 'consume liquid', 'consume alcohol').

Similar patterns of polysemy have been identified by other linguists (e.g. Apresjan 1974, Nunberg 1979, 1995, Bierwisch 1983, Pustejovsky 1995, Cruse 2000: 113), who claim that much – but not all – polysemy is “systematic”. This means that the relationships between senses are often shared by numerous polysemes and are sometimes found cross-linguistically (e.g. Murphy 2010: 89f., Dölling 2018: 1). Polysemy can therefore sometimes be recognised through the identification of one of these

patterns. One example of such a pattern is the PLANT/FOOD distinction, where one sense of a word describes a plant, and another describes the edible part of said plant (Dölling 2018: 2f.). In the sample, this can be seen with e.g. *Cranberry* and *Kartoffel* ‘potato’. Other patterns include: ANIMAL/FOOD (e.g. *chicken*) and TREE/WOOD (e.g. *oak*) (Dölling 2018: 5ff.). Examination of the possible relationship between senses is therefore a useful way of distinguishing polysemy from homonymy. However, this criterion can be subjective, since the perception of the relation between senses relies on personal and cultural interpretation.

A third way to distinguish polysemes from homonyms is to assess their morphosyntactic properties. In terms of German nouns, if the items have, for instance, distinct genders or plural markings, they are likely to be homonyms rather than different forms of a polyseme (e.g. Murphy 2010: 91, Mörth & Dressler 2014: 254). Mörth & Dressler (2014: 254) observe that there are some homonyms with the same gender but different plural forms, e.g. *die Bank (-en)* ‘bank [institution]’ vs. *die Bank (-e)* ‘bench’, some with different genders and plural forms, e.g. *der Leiter (-)* ‘leader’ vs. *die Leiter (-n)* ‘ladder’, and a few with different genders but the same plural form, e.g. *der Pony (-s)* ‘fringe’ vs. *das Pony (-s)* ‘pony’. However, this cannot always be used as a diagnostic tool, since not all homonyms have distinct morphosyntactic properties (e.g. Murphy 2010: 91).

In sum, none of the three main ways to differentiate between polysemy and homonymy is completely reliable in all cases. This, coupled with the fact that ambiguity tests are also not entirely consistent in their predictions, suggests that it may not always be possible to draw a distinction between cases of polysemy, homonymy and monosemy. Consequently, it has been proposed that homonymy, polysemy and monosemy are not discrete categories but instead form part of a continuum (Deane 1988: 327, 345, Tuggy 1993). Proponents of this view suggest that the continuum ranges from maximal distinctness (clear-cut homonymy) to maximal similarity (clear-cut monosemy, with clear-cut polysemy in the centre, as illustrated below.

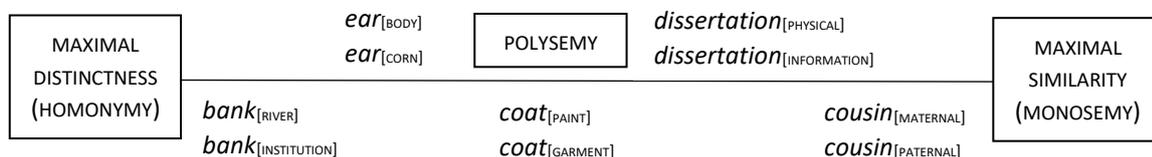


Figure 1: Depiction of the continuum of meaning variation as proposed by Tuggy (1993)

3.3.2.1.4 Polysemy vs. homonymy in the sample

The word forms in the sample which have been established as having multiple senses can be divided relatively easily into sets of homonyms and nouns which are polysemous using the three criteria mentioned above. There are no obvious cases of nouns which do not conform to either group.

Examples of homonyms

der Aviso (-s) 'small warship' vs. das Aviso (-s) 'notification'

These two lexical items are claimed to have entered German via different routes: *der Aviso* from Spanish *barca de aviso* 'boat used for conveying messages' via French *aviso*, and *das Aviso* from Old French *ce m'est a vis* 'that's my impression' via Italian *avviso* (Duden Online-Wörterbuch 2018a). Synchronically, the senses do not appear related. Additionally, the nouns belong to different genders, which further suggests that they are likely to be synchronically perceived as two separate nouns which are homonyms.

die Band (-s) 'music group'¹⁰ vs. das Band (-'er) 'ribbon' and der Band ('-e) 'book volume'

Die Band is an English loan which ultimately comes from French *bande*, and *das Band* originates from OHG *bant* which is historically linked to *der Band* (Pfeifer 2010, Kluge 2010: 29, Duden Online-Wörterbuch 2018b). From a synchronic perspective, there is no obvious semantic relation between the three forms. They also all have distinct genders and plural markers. Moreover, there is only a homographic – rather than homonymic – relation between *die Band* [bent] and *das* and *der Band* [bant].

die Leiter (-n) 'ladder'¹¹ vs. der Leiter (-) 'leader'

These are claimed to have divergent etymologies: *der Leiter* coming from OHG *leitari*, and *die Leiter* from OHG *leitara* (Pfeiffer 2010, Kluge 2010: 246, Duden Online Wörterbuch 2018e). There is no synchronic semantic relation between the senses and they have distinct genders and plural markings.

Examples of polysemes

The nouns discussed in the following paragraphs are analysed as being polysemous, given that their senses are claimed to derive from the same source, have the same morphosyntactic properties, and appear to be semantically related from a synchronic perspective.

¹⁰ *Die Band* as the head of *All-Star-Band* in the sample (see 3.3.3).

¹¹ *Die Leiter* as the head of *C-Dur-Tonleiter* in the sample (see 3.3.3).

As discussed in 3.3.2.1.3, the relationship between the distinct senses of *Cranberry* and *Kartoffel* fit the commonly-found pattern PLANT/FOOD, *Cranberry* referring to both the plant and the fruit, and *Kartoffel* referring to the plant and the tuber. The relationship between the senses of *Rotang* ('palm' and 'material') seems to conform to the broader pattern PLANT/MATERIAL, which would fit other polysemes such as *Baumwolle* ('plant' and 'fabric').

There are also nouns in the sample whose senses appear to be related by metaphorical extension, such as *Hirsch*, which can refer to a deer or be used as a profane term for a man. This would seem to conform to the more general pattern ANIMAL/HUMAN, which fits other nouns such as *Schwein* 'pig' or English *cow*. Other cases of figurative extension in the sample include: *Läufer* 'runner' and 'rotor', *Marke* 'brand/make' and 'peculiar person', and *Tamtam* 'musical instrument' and 'general noise/fuss'.

Now that a distinction has been made between the homonyms, polysemes and monosemes in the sample, it is important to consider how their categorisation might affect the analysis of these nouns in OGAT.

The analysis of monosemes is straightforward in terms of OGAT, since the input can only consist of one form and one meaning. Homonymy is also unproblematic, since it is widely agreed that the distinct senses belong to separate lexical entries (e.g. Cruse 2000: 97) and would thus serve as separate inputs to OGAT. The homonyms in the sample are discussed further in 3.3.4.5.

The main challenge for OGAT are the polysemes, since it is unclear which meaning should serve as the input. For example, *die Marke* in the sense of 'brand/make' is subject to the constraint $-E \rightarrow *M, *N$, meaning that feminine is correctly selected as optimal. However, *die Marke* in the sense of 'odd person' is subject to $-E \rightarrow *M, *N$ and $GEN.PERSON \rightarrow *F, *N$, which would incorrectly be assigned masculine in OGAT. It is therefore unclear whether these distinct senses should serve as separate inputs to gender assignment, or whether the input consists of all senses, none of the senses, or just one of the senses. Accordingly, the following section will discuss various approaches to polysemy.

3.3.2.1.5 Approaches to polysemy

The various approaches to polysemy suggested in the literature can be broadly categorised into two types. The first views polysemy as a semantic phenomenon, claiming that all senses of a polyseme are encoded in the lexicon, following the principle of lexical polysemy. The second type views polysemy as a pragmatic phenomenon, claiming that only one sense of a polyseme is stored in the lexicon and other senses are derived from the stored meaning in context, following the principle of lexical monosemy.

Lexical polysemy

The view that all senses of a polyseme are represented in the lexicon is labelled the “sense enumeration” strategy. According to this strategy, all conventionalised senses of a word are listed as either completely independent lexical entries or as distinct sub-entries of the same lexical entry (e.g. Cruse 2000: 97, Dölling 2018: 13). Comprehending a word therefore requires the selection of a sense from the available list (e.g. Pustejovsky 1995: 34ff., Dölling 2018: 13).

One approach which follows the sense enumeration strategy is that proposed in Katz & Fodor (1963), Katz & Postal (1964), Katz (1972) etc., where it is maintained that lexical senses consist of sets of necessary and sufficient conditions. According to this view, each sense of a polyseme is associated with its own set of conditions. These senses thus exist independently of one another. Homonyms are also claimed to be represented by separate sets of necessary and sufficient conditions, meaning that this approach does not make a distinction between polysemy and homonymy. In terms of OGAT, this view would dictate that each sense of a polyseme constitutes a separate input.

However, this approach has been criticised for various reasons. Firstly, it does not account for the fact that the distinct senses of a polyseme are related to each other (e.g. Pustejovsky 1995: 54, Asher 2011). Accordingly, it cannot explain cases where senses overlap (e.g. Pustejovsky 1995: 39) or any instances of systematic polysemy, where patterns of relations between senses are shared by multiple words and are often found cross-linguistically (e.g. Dölling 2018: 1). This ultimately leads to the consideration of polysemy as an arbitrary, accidental occurrence (e.g. Dobrić 2014: 152).

A further criticism of this approach comes in the form of psycholinguistic evidence. Falkum (2011: 63) states that of the studies conducted, the majority demonstrate a difference in the representation of homonymy and polysemy in the lexicon (e.g. Frazier & Rayner 1990; Williams 1992; Pickering & Frisson 2001; Klepousniotou 2002, 2007; Klepousniotou & Baum 2007; Klepousniotou et al. 2008; Beretta et al. 2005; Pykkänen et al. 2006; Brown 2008, cited in Falkum 2011: 63). Falkum (2011: 63) notes that very few of the studies report results which suggest that the representation of homonymy and polysemy is identical (e.g. Klein & Murphy 2001, 2002, cited in Falkum 2011: 63). The suggestion that all distinct senses of a polyseme are represented separately thus seems unfavourable.

Another approach which also follows the sense enumeration strategy is the network approach to polysemy (e.g. Lakoff 1987, Brugman 1988, Langacker 1988, Taylor 1995). Proponents of this approach maintain that all conventionalised senses of a polyseme are represented in the lexicon in a network formation, with the primary, prototypical meaning at the centre. This is similar to the previous approach in that all conventionalised senses are represented in the lexicon, but different in that it is able to account for the relatedness of senses, as well as allowing for a distinction between the

representation of polysemy and homonymy. The approach is also consistent with the substantial body of psychological evidence for typicality effects (e.g. Rosch 1973, 1975b, 1978; Rosch & Mervis 1975).

Whilst earlier versions of the model have been criticised for their lack of ability to explain cases of systematic polysemy and to predict new senses based on existing patterns of relations between senses (e.g. Murphy 2010: 104, Dölling 2018: 14), more recent models following the network approach (e.g. Tyler & Evans 2001, 2003) are argued to overcome many of these problems (see e.g. Falkum 2011: 42).

In terms of polysemes in the sample, it might be argued that – in line with the network approach – it is the primary, prototypical sense at the centre of the network that determines which constraints apply to a polyseme and, consequently, its gender. In order to identify this prototypical sense, a number of criteria have been proposed (e.g. Rice 1996: 145f.; Tyler & Evans 2001, 2003: 45-50; Evans 2004: 97, 2005), which state that the prototypical sense is likely to be the sense which: is most frequent, is historically earliest attested, is used in composite forms, has a plausible cognitive antecedent, relates to human experience, and is the sense from which other senses can be derived.

Lexical monosemy

In the monosemic approach, or “one representation hypothesis”, it is claimed that for each polyseme, there is a single lexically-stored representation from which all senses are contextually derived (e.g. Caramazza & Grober 1976, Allerton 1979, Nunberg 1979, Ruhl 1989, Pustejovsky 1995). This approach thus argues for a distinction between the semantic part of a lexical item (the stored representation), and the pragmatic, contextual part (the derived sense(s)) (Lewandowska-Tomaszczyk 2007: 152f.).

Among proponents of lexical monosemy, there is debate as to the nature of the single, lexically-stored representation (e.g. Falkum & Vicente 2015: 5-8). Proposals range from a basic, abstract representation (e.g. the representation as an abstract meaning shared by all senses (Ruhl 1989); the representation as constraints on what the word may/may not express (Carston 2012); the representation as underspecified and only enriched if the context requires (Frisson 2009)) to one which is informationally rich (e.g. the representation providing considerable informational content of which hearers select only the relevant part (Pustejovsky 1995)). It has also been proposed that the stored representation is the sense of the word that is most conventional (Nunberg 1979).

As with the network approach, the monosemic approach is able to account for the fact that the senses of a polyseme are related, and for the differences between polysemy and homonymy. The approach is also the most parsimonious of those discussed, since it requires minimal semantic representation in the lexicon. Furthermore, proponents of this view provide various explanations for systematic

polysemy, such as the claim that extended senses are derived from the lexically-stored representation via lexical rules, which are said to usually apply to multiple words, thus accounting for systematicity in the lexicon (e.g. Copestake & Briscoe 1995).

In terms of OGAT, then, it could be argued that – according to the monosemic approach – it is the single representation in the lexicon which determines the relevance of semantic constraints to the polyseme. However, the practical application of this idea to OGAT is problematic, since it is unclear what the single representation consists of. Indeed, Dölling (2018: 16) criticises this approach for its lack of reliable principles to determine what the representation comprises.

The monosemic approach is also criticised for its inability to account for irregular cases of polysemy where the relationship between a word’s senses is not shared by other polysemes (e.g. Murphy 2010: 101). For instance, Pustejovsky (1995) treats cases of irregular polysemy in the same way as homonyms, i.e. each sense belonging to a separate lexical entry, and therefore cannot account for the relatedness between the senses.

To summarise, it seems that the most useful approach for the present study out of those discussed is the network approach to polysemy, which would dictate that the primary, prototypical meaning of a polyseme is used as the input to gender assignment. Unlike the monosemic approach, the network approach provides a clear idea of what the centrally-represented item for each polyseme consists of, as well as clear strategies for determining it.

3.3.2.1.6 Polysemy in the sample

For many polysemes in the sample, how polysemy is lexically represented is unimportant, since all possible cases yield the same outcome in OGAT. That is to say, regardless of whether the senses constitute separate inputs, or whether the input consists of all senses or just the primary sense, OGAT produces the same result. This is illustrated by the examples below. The senses listed have been taken from the Duden.

Noun	Senses	Applicable constraints	OGAT outcome
<i>der Läufer</i>	• runner [sport]	-ER SUFF. →*F,*N; GEN. PERSON →*F,*N	✓ (M)
	• rug; bishop [chess]; rotor	-ER SUFF. →*F,*N	
<i>der Hirsch</i>	• deer; red deer	-(C)(C)// MONO →*F,*N; CVCC →*F; DOM. ANIMALS →*F; MAMMALS →*F	✓ (M)
	• male red deer	-(C)(C)// MONO →*F,*N; CVCC →*F; DOM. ANIMALS →*F; MAMMALS →*F; MALE ANIMALS →*F,*N	
	• [Berlin, colloquial] expert, dab hand	-(C)(C)// MONO →*F,*N; CVCC →*F; GEN. PERSON →*F,*N	

	<ul style="list-style-type: none"> [humorous] profanity for a man [humorous] cuckold 	-(c)(c)// MONO →*F,*N; CVCC→*F; MEN→*F,*N	
<i>das Motiv</i>	<ul style="list-style-type: none"> motive, motivation 	-IV→*M,*F	✓ (N)
	<ul style="list-style-type: none"> general theme, picture or form of a work of literature or art; artistic composition 	-IV→*M,*F; PICTURES→*M	
	<ul style="list-style-type: none"> [Music] recognisable part of a melody, characteristic of a certain composition 	MUS. TERM/COMP→*M; -IV→*M,*F	
<i>die Klasse</i>	<ul style="list-style-type: none"> school year; classroom; university faculty [Biology] organisms with similar features group of vehicles/boats quality level; part of an ordinal scale 	-E→*M,*N	✓ (F)
	<ul style="list-style-type: none"> group of pupils [Sociology] group within population [Sport] group of sportspeople/teams 	-E→*M,*N; GROUPS→*M,*N	

However, there are at least three cases in the sample for which determining the input is crucial, since the various possibilities cause OGAT to yield different outcomes.

Noun	Senses	Applicable constraints	OGAT outcome
<i>die Marke</i>	<ul style="list-style-type: none"> brand, make, trademark identification tag, stamp specific point, marking 	-E→*M,*N	✓ (F)
	<ul style="list-style-type: none"> [slang] peculiar person 	-E→*M,*N; GEN. PERSON→*F,*N	x (M)
<i>die Perle</i>	<ul style="list-style-type: none"> shiny bead from oysters bead; bubble or droplet [hunting] bump on antlers 	-E→*M,*N	✓ (F)
	<ul style="list-style-type: none"> [colloquial, humorous] housemaid [archaic slang] faithful girlfriend 	-E→*M,*N; WOMEN→*M,*N	✓ (F)
	<ul style="list-style-type: none"> [colloquial, humorous] someone who provides valuable services 	GEN. PERSON→*F,*N; -E→*M,*N	x (M)
<i>das Tamtam</i>	<ul style="list-style-type: none"> musical instrument (type of gong) 	MUS. INSTR.→*M; SEM. EQUIV.→*F	✓ (N)
	<ul style="list-style-type: none"> [colloquial] fuss, racket 	SEM. EQUIV.→*F,*N	x (M)

In line with the network approach, the primary, prototypical senses of these nouns must be established, as this is likely to be the sense which is relevant to gender assignment. This can be done using some of the criteria outlined in 3.3.2.1.5, such as: the sense most frequently used, the historically earliest attested sense, the sense used in composite forms and the sense from which other senses can be plausibly derived.

To determine the most frequently-used sense, the 50 most recent entries for each of the nouns in the DWDS-Kernkorpus were examined to assess which sense is used in each case. For *Perle*, the sense 'shiny bead from oysters' was used in 49 cases, for *Marke*, the sense 'brand/make' was used in 50 cases, and for *Tamtam*, the sense 'musical instrument' was used in 37 cases.

These same three senses also appear to be the senses from which the others can be plausibly derived, as figurative extensions of the literal sense. For *Perle* and *Marke*, it is also these senses which are used

in composite forms, e.g. *Markenname*, *Automarke*, *Markengesetz*, *Markeninhaber* and *Perlenkette*, *Perlenohrring*, *Perlenauster*, *Perlenfarm* etc.

Finally, the same results are also produced when examining the earliest attested sense. In each case, the concrete sense is attested first, with the metaphorical sense(s) emerging later. For instance, Pfeifer (2010) states that *Perle* exists in OHG as *perala/perula*, and that the figurative sense denoting something ‘schön, edel, wertvoll’ emerges only in the 16th century. Similarly, *Tamtam* is said to have been borrowed from French *tamtam* ‘gong/drum’ in the early 19th century, and only in the late 19th century did it begin to be used to denote any drum-like sound and, subsequently, to mean ‘a loud fuss/racket’ (Pfeifer 2010).

The criteria therefore all point towards the same senses (highlighted in bold in the table) being the prototypical senses for each noun. The results are consistent with the fact that it is with these prototypical senses that OGAT is able to assign the correct gender to each noun, as demonstrated above.

3.3.2.2 Form variation

Another group of nouns in the sample of which there are multiple variants that are relevant to gender assignment are those with variation in their form. Specifically, there are three nouns in the sample which have more than one possible realisation in Standard German according to the Duden. These are: *Jazz*, *Islam* and *Avis*.

Noun	Possible realisations
<i>der Jazz</i>	[dʒæz], [dʒɛs], [jats]
<i>der Islam</i>	[ɪs'la:m], ['ɪslam]
<i>der/das Avis</i>	[a'vi:], [a'vi:s]

All three of these nouns are loanwords and therefore this variation is unsurprising, given that variation in the realisation of loans is a common phenomenon (e.g. Poplack et al. 1988, Russ 1994, Muhvić-Dimanovski 1995, Greisbach 2003, Lev-Ari et al. 2014). It has been observed that variation in the realisation of loanwords could be due to a number of factors, including: the speaker’s regional background (e.g. Russ 1994: 55, Muhvić-Dimanovski 1995), the speaker’s age and socio-economic background (e.g. Greisbach 2003), the speaker’s knowledge of the donor language (e.g. Poplack et al. 1988), and the speaker’s perception of the donor language (Lev-Ari et al. 2014). It has also been claimed that there may be intra-speaker variation, as it has been suggested that the speaker’s

pronunciation will vary according to their interlocutor’s proficiency in the donor language (Lev-Ari et al. 2014).

Variation in the form of a noun has the potential to be problematic for OGAT, as it could result in the prediction of different genders for different forms of the same noun. However, unlike with polysemy, this does not occur in the sample. OGAT predicts the same, correct gender for all variants of the three nouns. For *Avis*, this is because both variants are subject to exactly the same constraints. For *Jazz* and *Islam*, the variants are subject to some different constraints, but the OGAT outcome is the same, as demonstrated below. Whether this issue would be problematic for any nouns outside of the sample is a matter for further investigation.

Noun	Realisation	Applicable constraints	OGAT outcome
<i>der Jazz</i>	[dʒæz]	CCVC→*F,*N MUSIC TYPES→*F,*N	✓ (M)
	[dʒɛs]	CCVC→*F,*N MUSIC TYPES→*F,*N	
	[jats]	CVCC→*F MUSIC TYPES→*F,*N -C S MONO→*F,*N V _{SHORT} STOPC→*F,*N	
<i>der Islam</i>	[ɪs'la:m]	RELIGIONS→*F,*N ¹² V _{STRESSED} [m]→*M,*F	✓ (M)
	['ɪslam]	RELIGIONS→*F,*N	
<i>der/das Avis</i> ¹³	[a'vi:]	SEM. EQUIV.→*F,*N (DER HINWEIS) /SEM. EQUIV.→*M,*F (DAS AVISO)	✓ (M/N)
	[a'vi:s]	SEM. EQUIV.→*F,*N (DER HINWEIS) /SEM. EQUIV.→*M,*F (DAS AVISO)	

3.3.3 Complex nouns and “strong rules”

Another subset of nouns which require further investigation in order to ascertain whether OGAT is able to account for their genders are complex nouns such as compounds (see 1.3.1.4) and nouns derived via affixation (see 1.3.1.1). Related to this is the question raised in 1.4.5.3 of how OGAT might be able to account for Steinmetz’s (1986: 194) observation that some GENDER FEATURES constraints – mostly those involving suffixes – are “strong”, in that they override all others. For instance, nouns with the suffix *-ung* are assigned feminine regardless of any other features of the noun (Steinmetz 1986: 194).

¹² See 3.4.1 for details of this constraint.

¹³ See Section 3.3.4.8 for further discussion of *Avis*.

As stated in 1.4.5.3, Rice (2006: 1409) claims that the genders of compounds and affixed nouns as well as Steinmetz's (1986) observation regarding "strong" constraints can be accounted for in OGAT through reference to headedness:

Regardless of the gender-relevant features on a compound, it will be assigned the gender of its rightmost element. In an OT conceptualization, this means that the force of the rightmost element outranks any other relevant constraints. This will also be true of some morphemes which Steinmetz (1986) calls strong endings; these morphemes – such as the neuter assigning *-chen* diminutive suffix in German – will also outrank any other constraints sensitive to gender-relevant features. This is compatible with our proposal about crucial equal ranking, since these cases presumably do not require explicit reference to language specific material, but rather refer to a morpheme's status as a head. (Rice 2006: 1409)

Accordingly, the notion of headedness will be explored below (3.3.3.1), followed by a discussion of how OGAT can account for the genders of multi-word units (3.3.3.2) and nouns derived via affixation (3.3.3.3) in the sample.

3.3.3.1 Heads

The notion of a head in the morphological description of complex words is an extension of the notion of a phrasal head in syntactic theory (see e.g. Corbett et al. 1993). The term has been adopted by lexical morphologists (e.g. Lieber 1980, Williams 1981, Selkirk 1982, Scalise 1988) to describe the element in complex words which typically determines certain properties of the word as a whole, e.g. gender. It is claimed that such features are transmitted from the head to the "mother node" (which dominates the entire complex noun) via "feature percolation" (e.g. Lieber 1981, 1989; Williams 1981 etc.). Due to their proposed relevance to gender assignment, it is important to consider how the heads of complex nouns may be identified.

3.3.3.1.1 Head identification

There are a variety of approaches to determining heads which can be broadly categorised into four types: semantic, morphosyntactic, categorial and morphological.

The semantic approach to head determination claims that the head is the constituent which determines the semantic features of a complex word. For instance, *man* can be argued to be the head of *postman*, since it is the element which transmits the features [+ANIMATE] and [+HUMAN] to the whole complex noun, as in (34).

(34) $[_{[+ANIMATE, +HUMAN]}[_{-ANIMATE, -HUMAN}]post]_{[_{[+ANIMATE, +HUMAN]}man]}$

In terms of identifying the head on a semantic level, it is claimed that there is typically a subset relationship between the head and the complex word as a whole, whereby the complex word denotes a subset of the head. This results in the meaning “a kind of X” where X is the head, e.g. *warship* is a kind of *ship* (e.g. Allen 1978: 105-108, Hoeksema 1992: 120, Bauer 2017: 37). Bauer (1990: 5, 2017: 37) labels this the “hyponymy criterion”.

The morphosyntactic approach to head determination claims that the head of a complex word is the morphosyntactic locus, i.e. the element which bears inflectional markers which belong to the entire construction (e.g. Zwicky 1985, Hudson 1987, Namiki 2001). For example, *Arzt* might be considered the head of *Tierarzt*, since it is the element to which plural and genitive markers attach: *die Tierärzte* – **die Tierearzt, des Tierarztes* – **des Tieresarzt*.

The categorial approach to head identification states that the head is the constituent which determines the syntactic category of the word (e.g. Williams 1981; Kiparsky 1982: 133; Gagné & Spalding 2006: 160; Kageyama 2008, 2009; Lefebvre 2014: 168). According to this view, nominal compounds are nouns because their head is a noun, e.g. [_N[_v*Fahr*][_N*gast*]].

Finally, the morphological approach to head determination defines the head as the element which specifies morphological features such as gender, number and inflectional class etc., but is not necessarily the locus of inflection (e.g. Scalise et al. 2009, Scalise & Fábregas 2010). For instance, *Flasche* can be seen as the gender-determining morphological head of [_{FEM}[_{NEUT}*Bier*][_{FEM}*flasche*]], and *Bier* can be seen as the gender-determining morphological head of [_{NEUT}[_{FEM}*Flasche(n)*][_{NEUT}*bier*]].

Prototypically, a head of a complex word can be considered a head using all four of these approaches (e.g. Dressler 2006: 31). However, for some complex nouns, not all approaches lead to the same element being identified as the head (e.g. Di Sciullo & Williams 1987: 26, Rainer 1993: 57ff., Dressler 2006: 32f.). Consequently, some theories of headedness allow for certain constituents to be classified as heads for some features and non-heads for others, e.g. the Relativised Head Hypothesis (Di Sciullo & Williams 1987) and the Three-Level Head Hypothesis (Scalise et al. 2009, Scalise & Fábregas 2010). All four approaches will be considered when determining the heads of the complex nouns in the sample, with a particular focus on the morphological head, since this is the element which is claimed to determine gender.

3.3.3.1.2 Head position

A proposal which has been hugely influential on theories of German gender assignment is the idea that the head of a complex word always appears in the same position within the construction.

Specifically, Williams (1981: 248) claims that ‘in morphology, we define the head of a morphologically complex word to be the righthand member of that word.’ This is known as the Right-hand Head Rule (RHR) and is the foundation for Köpcke & Zubin’s (1983: 9, 1984: 44f., 1997: 28) Last Member Principle (LMP) (see 1.3.1.4).

The RHR has, however, received much criticism for being overly simplistic and cross-linguistically untenable (e.g. Lieber 1980, Bauer 1990), since not only are there languages with left-headed structures (e.g. Zwanenburg 1992, Scalise 1992: 179ff., Rainer 1993: 57, Semenza & Mondini 2006, Levy et al. 2006), but there is also evidence of head position varying within a language, i.e. casting doubt on the possibility of a head position parameter (e.g. Hoeksema 1992: 130, Scalise & Fábregas 2010: 53). Therefore, while the RHR may be upheld as a universal tendency (Dressler 2006: 33, Scalise et al. 2009: 50), it is not without exception.

Nonetheless, German largely conforms to this universal tendency, with a high proportion of structures having an element which is considered the head from a semantic, morphosyntactic, categorial and morphological perspective in the right-hand position. This has led to the conclusion that the canonical position of the head in German is on the right (e.g. Guevara & Scalise 2009: 114), which explains the broad acceptance of the LMP.

The following sections will consider whether OGAT is able to account for the multi-word units and affixed nouns in the sample using the notion of headedness, as Rice (2006: 1409) proposes.

3.3.3.2 Multi-word units and OGAT

Approximately half of the nouns in the sample are so-called “multi-word units” (e.g. Plag 2003). Most of these are compounds and three are lexicalised phrases. Note that compound loanwords (3.3.1) and proper nouns (3.3.4) are discussed separately.

3.3.3.2.1 Compounds

The vast majority of compounds in the sample are those traditionally classified as determinative endocentric compounds, the most common type of compound in German (e.g. Hüning 2008: 197). In the sample, these all have prototypical heads in the canonical position for German, i.e. on the right. The heads of these compounds conform on all four levels of headedness, i.e. semantically, morphosyntactically, categorially and morphologically. For example, *Baum* ‘tree’ is the head of *Mandelbaum* ‘almond tree’, which is evident from a semantic perspective, in that the compound denotes a kind of tree; from a morphosyntactic perspective, in that *Baum* carries inflection (e.g. *die Mandelbäume* vs. **die Mandelbaum*); from a categorial perspective, in that *Baum* shares its syntactic

category with the compound; and from a morphological perspective, in that *Baum* specifies the gender of the compound: [_{MASC}[_{FEM}*Mandel*]][_{MASC}*baum*]]. Other compounds in the sample of this type include: *Hubbrücke, Erschöpfungssyndrom, Fraktionssprecher, Kontaktformular* etc.

Also in the sample are a small number of compounds usually classified as exocentric, e.g. *Pimpernuss* ‘bladder nut [plant]’. Traditionally, these are thought of as not containing an element conforming to the semantic idea of headedness, however, a number of linguists (e.g. Booij 1992: 39, Bauer 2008: 59) argue that a hyponymous relationship could exist between *Pimpernuss* and *Nuss* ‘nut’, if *Nuss* is interpreted as a metonymical/synecdochic extension of the referent of the compound. *Nuss* – appearing in the canonical right-hand position – can also be considered the head from a morphosyntactic, categorial and morphological perspective. This is because *Nuss* is the locus of inflection: *Pimpernüsse*, the category determinant: [_N[_V*Pimper*]][_N*nuss*]], and the gender and plural determinant: [_{FEM}[*Pimper*]][_{FEM}*nuss*]], *Pimpernüsse*.

3.3.3.2.2 Lexicalised Phrases

There are also three complex structures in the sample which are considered lexicalised phrases rather than compounds: *Perpetuum mobile, Nomen proprium* and *Gesellschaft für Sport und Technik (GST)*.¹⁴ For *Nomen proprium* and *Perpetuum mobile*, this is evidenced by the fact that agreement is displayed on all elements in the construction (including the modifying adjective) rather than just on the head: *die Nomina propria, die Perpetua mobilia* (cf. *die Hochschulen* vs. *die hohen Schulen* (Dressler 2006: 27)) (e.g. Plag 2003: 173f., Booij 2005: 82). For *Gesellschaft für Sport und Technik*, it is evidenced by the head having a PP as a post-modifier, rather than as a pre-modifier as is required for compounds containing syntactic phrases (e.g. Plag 2003: 174). The phrasal status of these constructions is also reflected in their orthography; the constituents are written separately rather than together as is typical for compounds (Hüning 2008: 200, Schlücker & Hüning 2009: 212, Schlücker & Plag 2011: 1541). The proposed structures are therefore as follows:

[_{NP}[_N*Perpetuum*]] [_{ADJP}[_A*mobile*]]]

[_{NP}[_N*Nomen*]] [_{ADJP}[_A*proprium*]]]

[_{NP}[_N*Gesellschaft*]] [_{PP}[_P*für*]] [_{NP}*Sport und Technik*]]]

In terms of the heads of these structures, it seems clear that *Perpetuum, Nomen* and *Gesellschaft* all have head status. Semantically, the structures refer to e.g. a kind of *Nomen* or a kind of *Gesellschaft*.

¹⁴ On the basis that the gender of an initialism (or any shortened form) is based on the gender of the full form (see 1.3.1.5).

Morphosyntactically, *Gesellschaft* is the locus of inflection (e.g. *die Gesellschaften für Sport und Technik*). Categorially, *Perpetuum* and *Nomen* determine the category of the phrase, and morphologically, these elements specify e.g. the gender of the constructions.

In each case, the heads are the leftmost element of the construction, rather than the rightmost element as is typical for compounds. Lexicalised phrases are, therefore, not bound by the RHR or the LMP (see also 1.3.1.5). While *Perpetuum mobile* and *Nomen proprium* are phrases borrowed from Latin whose structure could be considered atypical for that reason, this argument can be upheld on the basis of *Gesellschaft für Sport und Technik*.

3.3.3.2.3 Multi-word units and OGAT

Above it has been demonstrated that multi-word units typically contain a head noun which is considered a head according to most, if not all, criteria. For compounds, this is typically the rightmost noun, in line with the RHR/LMP. For lexicalised phrases, however, the head can appear in various positions, meaning that the RHR/LMP is not valid for these structures.

Since it is the head noun in these structures whose gender is inherited by the unit as a whole, any model of gender assignment should focus exclusively on the properties of the head, disregarding any properties which are only relevant to the unit as a whole. This is in line with Zubin & Köpcke's (1984: 85) statement that gender is assigned 'with no regard to the semantic content of the composite [as a whole]'

Turning to the multi-word units in the sample, this does indeed seem to be the case. For structures such as those in the table below, the correct gender prediction can only be made on the basis of constraints which apply to the head noun, rather than those which apply only to the structure as a whole.

Multi-word unit	Constraints	OGAT prediction	Head noun	Constraints	OGAT prediction
<i>der Affenbrotbaum</i>	TREES → *M, *N	x (F)	<i>Baum</i>	-N(C)(C) MONO → *F, *N VV _{DIPHTH,C} MONO → *F CVC MONO → *F	✓
<i>der Handball</i>	GAMES → *M, *F	x (N)	<i>Ball</i>	-L MONO → *F CVC MONO → *F	✓
<i>die Janitscharenmusik</i>	MUSIC TYPES → *F, *N	x (M)	<i>Musik</i>	-IK → *M, *N ACAD.DISCIPLINES → *M, *N	✓
<i>das Nuteisen</i>	STICK → *N	x (M)	<i>Eisen</i>	METALS → *M, *F ELEMENTS → *M, *F -EN → *F	✓

However, there are also cases in the sample in which the correct gender could be predicted on the basis of constraints which apply either to the whole unit or just the head noun.

Multi-word unit	Constraints	OGAT prediction	Head noun	Constraints	OGAT prediction
<i>die GST = Gesellschaft für Sport und Technik</i>	GROUPS → *M, *N	✓	<i>Gesellschaft</i>	-SCHAFT → *M, *N GE- → *M, *F GROUPS → *M, *N	✓
<i>das Nomen proprium</i>	WORD CLASSES → *M, *F	✓	<i>Nomen</i>	WORD CLASSES → *M, *F -EN → *F	✓

The one exception to this is *Juchtenleder*, for which OGAT is only able to make the correct prediction based on the properties of the compound as a whole. This is because OGAT is unable to predict that *Leder* is a neuter noun (see 3.5.2 for further exceptions).

Multi-word unit	Constraints	OGAT prediction	Head noun	Constraints	OGAT prediction
<i>das Juchtenleder</i>	LEATHER TYPES → *M, *F	✓	<i>Leder</i>	-ER → *F, *N	x (M)

As far as multi-word units are concerned, then, the focus of OGAT should be on the head noun, as Rice (2006: 1409) suggests. In terms of accounting for this in the model, there are a number of possibilities. It could be the case, for instance, that only the head noun of a multi-word unit serves as the input to OGAT. Alternatively, it could be the case that the whole multi-word unit enters OGAT, but all constraints are specified as applying only to the head noun. In either case, the assignment of gender to multi-word units can be categorised as a “head operation” (Hoeksema 1986), since it is a morphological operation which applies only to the head noun.

3.3.3.3 Nouns derived via affixation and OGAT

Given the relative success of the headedness approach for German multi-word units, it is unsurprising that the notion of a structural head has been extended to the analysis of nouns derived via affixation (e.g. Toman 2001: 233). For instance, the suffix *-er* in a complex noun such as *Lehrer* might be interpreted as a kind of bound noun which acts as the head, since it fulfils many of the relevant criteria for headedness (e.g. Toman 2001: 233). In *Lehr_V-er_{N:MASC}* the suffix *-er* determines the syntactic category and the masculine gender of the noun, it is the element to which inflectional markers attach, and, in terms of the semantic criterion, the resulting derivative could be said to denote a kind of agent, i.e. a kind of noun in *-er* (Bauer 1990: 5). It therefore might be argued that the suffix *-er* is the head of the structure from which features such as gender are inherited by the word as a whole.

This line of reasoning supports Rice’s (2006: 1409) claim that the genders of derived nouns can be explained in the same way as the genders of compounds, i.e. without reference to language-specific material and therefore without implementing ranking in OGAT, but rather by referring to the head status of the gender-determining element. Crucially, this would enable OGAT to account for

Steinmetz's (1986) proposal of "strong rules" (see 1.4.5.3), as OGAT would focus solely on the properties of the head, thus rendering any gender-relevant features of the structure as a whole irrelevant.

However, for a number of reasons, many linguists have expressed scepticism about extending the notion of headedness to derived nouns (e.g. Reis 1983, Bauer 1990, Beard 2001, Haspelmath & Sims 2010: 149).

Firstly, the head status of many German affixes is unclear. For instance, while the suffixes *-heit/keit*, *-schaft* and *-ung* have the ability to determine the syntactic category and gender of a noun, and are the elements to which inflection attaches, they fail to meet the semantic criterion for headedness. Haspelmath & Sims (2010: 149), for example, reject the idea that suffixes which derive abstract nouns fulfil the hyponymy criterion, stating that, e.g. *reality* cannot plausibly be defined as a kind of *-ity* or, for that matter, a kind of *real*. Indeed, the argument that a noun might denote e.g. "a kind of *-ung*" is weakened by the fact that many of these suffixes are 'semantisch offen [semantically open]' (Fleischer & Barz 2012: 225), in that there is a considerable amount of variation in terms of the semantic effect they have on the base, cf. *Wohnung* 'flat', *Meinung* 'opinion', *Forschung* 'research' (e.g. Hamm & Kamp 2009: 6f.).

Similarly, other suffixes, such as *-chen*, *-lein*, *-in* and *-ling* might be argued to be heads since they are the locus of inflection and determine gender (e.g. Kastovsky 2006: 2336). However, in terms of the categorial criterion, they are unable to determine word class since they are class maintaining, meaning that it is the base which determines the syntactic category of the word. Additionally, from a semantic perspective, Bauer (1990: 6) argues that in nouns derived using class-maintaining suffixes, the base is more likely to be the head than the suffix. For the English derivative *duckling*, for instance, he maintains that the noun as a whole denotes a kind of *duck*, rather than a kind of small thing, thus making *duck* the semantic head.

The headhood of the suffixes *-tum*, *-nis* and *-sal* is even more doubtful, especially in terms of gender. While they are able to determine the syntactic category¹⁵ and plural form of a resulting derivative, they do not determine the same gender for all derivatives. As stated in 1.3.1.1, derivatives in *-nis* and *-sal* may be feminine or neuter, and those in *-tum* may be neuter or masculine. The analysis of these suffixes as heads is therefore incompatible with Rice's (2006: 1409) claim that headedness accounts

¹⁵ *-tum* is now only productively used as a class-maintaining suffix, but there are nevertheless instances of it determining word class, e.g. *N [ADJ Heilig] Ntum*.

for “strong rules”. Additionally, the head status of these suffixes is questionable from a semantic perspective, for the same reasons as those listed above for the abstract suffixes.

It has also been argued that certain prefixes such as *Ge-* may have head status (e.g. Mills 1984, Lieber 1992). This is because it is able to determine the syntactic category, as well as often determining neuter gender and a collective meaning of the noun (e.g. Olsen 1991: 333). However, it does not meet all of the criteria. *Ge-* does not fulfil the morphosyntactic criterion, since it is not the locus of inflection; inflectional markers attach instead to the base, e.g. *die Gespann+e_{PL}, des Getier+S_{GEN}*. There are also cases of non-neuter nouns with an initial *Ge-* in which the *Ge-* is plausibly analysed as a prefix, e.g. *die Gezeit* and *die Gefahr*. Additionally, it does not occur in the canonical (right-hand) position for German, thus violating the RHR/LMP.¹⁶

Crucially, Booij (2000: 861), argues that, even if the notion of headedness were extended to all derived nouns, the inheritance of gender by a derived noun from an affix would not be ‘inheritance strictu sensu because the affix only exhibits these properties [e.g. gender specification] within a complex word, whereas in the case of compounds there is independent motivation for the morpho-syntactic properties that we assign the head, since this head also occurs as an independent word’. That is to say, while the head noun in a multi-word unit has a gender as an independent entity, an affix does not, since it only has gender-relevant properties within the context of a whole derived noun.

Therefore, it would be nonsensical for OGAT to operate on the basis of an affix alone, as is necessary with head nouns in multi-word units. That is to say, an affix could not alone serve as the input to OGAT, and all GENDER FEATURES constraints could not be specified as applying only to an affix. If that were the case, then an affix would be treated in the same way as a monosyllabic noun and subject to e.g. various phonological principles, which are of course irrelevant to affixes. Therefore, in terms of gender assignment, it is important to consider a derived noun as a whole, as that is the context in which an affix has a gender-determining role.

Finally, as stated in 1.4.5.3, Rice’s (2006: 1409) claim that “strong rules” can be accounted for by referring to all suffixes as heads, if true, would only explain “strong” morphological rules and not any

¹⁶ It has, however, been argued by Plank (1986) that *Ge-* is not a prefix but a circumfix *Ge_e*, and that the final *-e* is not always realised. Plank (1986) claims that the presence of final *-e* is not always visible but still often apparent through a stem vowel change in the base, e.g. *das Gehölz*. It could therefore be argued that the rightmost element is the final *-e* of the circumfix *Ge_e* (even if covert), which acts as the head. However, arguments against this analysis include the fact that there are nouns with initial *Ge-* which do not have final *-e* or a stem vowel change, e.g. *das Gesuch, das Gespann*, making the presence of a covert final *-e* seem implausible (and furthermore, the DWB does not contain any evidence to suggest that a final *-e* was present historically on these nouns). Additionally, even if Plank’s (1986) analysis were valid, Bauer (2017: 32) maintains that circumfixes as heads would still violate the RHR, since they occur on both the left and right-hand side of a base.

cases of “strong” semantic or phonological constraints, such as those found to have 100% coverage in 2.3, e.g. CHEESES→*F,*N; LANGUAGES→*M,*F; METALS→*M,*F; LETTERS/MUS.NOTES→*M,*F; VCCC→*M. These cannot, of course, be accounted for by headedness.

In sum, it has been demonstrated that headedness is able to account for gender assignment to multi-word units in the sample, but that its extension to derived nouns is unconvincing, and, moreover cannot fully account for Steinmetz’s observation regarding “strong” constraints in German. Consequently, “strong” constraints must be accounted for in OGAT in some other way.

It is proposed that such constraints can instead be accounted for in OGAT most appropriately by means of ranking. Specifically, it is proposed that all “strong” constraints are ranked in a block above other GENDER FEATURES constraints, as in (35). This enables OGAT to account for the fact that, when a “strong” constraint applies to a noun, it is this constraint that ultimately determines the gender of the noun, regardless of any other gender-relevant features the noun may have. Note that the relationship between the constraints within this block, particularly the interaction between “strong” semantic and form-based constraints, will be explored further in 3.3.4 and Chapter 4.

(35)

	GENDER FEATURES			DEFAULT HIERARCHY		
	“STRONG” CONSTRAINTS	OTHERS	SEM. EQUIV. ¹⁷	*N	*F	*M
	E.G. -LING→*F,*N; ¹⁸ -CHEN→*M,*F; LANGUAGES→*M,*F	E.G. TREES→*M,*N; ROCKS/MINERALS→*F,*N; V _{STRESSED} [m]→*M,*F	SEM. EQUIV. →*M/*F/*N			
<i>der</i> _____						*
<i>die</i> _____					*	
<i>das</i> _____				*		

Such ranking of “strong” constraints means that OGAT is able to account for a number of nouns in the sample which it otherwise could not if all GENDER FEATURES constraints were equally ranked, e.g. *das Hexaeder* and *das J* as shown in (36) and (37).

(36a)

<i>Hexaeder</i> 'hexahedron'	GENDER FEATURES		DEFAULT HIERARCHY		
	-EDER→*M,*F	-ER→*F,*N	*N	*F	*M
☞ <i>der Hexaeder</i>	*				*
<i>die Hexaeder</i>	*	*!		*	
<i>das Hexaeder</i>		*	*!		

¹⁷ See 3.3.1.2.

¹⁸ Note that the fact that suffixes are only gender-relevant features when they appear word-finally can be accounted for by specifying this in the constraint definition or, formally, by the insertion of a word boundary, e.g. -LING#→*F,*N “nouns with the final suffix *-ling* cannot be feminine or neuter”.

(36b)

<i>Hexaeder</i> 'hexahedron'	GENDER FEATURES		DEFAULT HIERARCHY		
	"STRONG" CONSTRAINTS	OTHERS	*N	*F	*M
	-EDER→*M,*F	-ER→*F,*N			
<i>der Hexaeder</i>	*!				*
<i>die Hexaeder</i>	*!	*		*	
<i>☞ das Hexaeder</i>		*	*		

(37b)

<i>J</i> [jɔt] 'J'	GENDER FEATURES			DEFAULT HIERARCHY		
	LETTERS→*M,*F	V _{SHORT STOP} C→*F,*N	CVC→*F	*N	*F	*M
<i>☞ der J</i>	*					*
<i>die J</i>	*	*	*!		*	
<i>das J</i>		*		*!		

(37b)

<i>J</i> [jɔt] 'J'	GENDER FEATURES			DEFAULT HIERARCHY		
	"STRONG" CONSTRAINTS	OTHERS		*N	*F	*M
	LETTERS→*M,*F	V _{SHORT STOP} C→*F,*N	CVC→*F			
<i>der J</i>	*!					*
<i>die J</i>	*!	*	*		*	
<i>☞ das J</i>		*		*		

The ranking of "strong" constraints above other GENDER FEATURES constraints will be tested in Chapter 4.

3.3.4 Subset of 55 nouns

A further group of nouns from the sample which require further investigation in order to establish whether OGAT is able to account for their genders is the subset of 55 nouns which, as stated in 2.2, have been identified as being exceptional in some way. These include: placenames (3.3.4.1), personal names (3.3.4.2), plural nouns (3.3.4.3), regional nouns (3.3.4.4), archaic nouns (3.3.4.5), homonyms (3.3.4.6), adjectival nouns (3.3.4.7) and multiple-gender nouns (3.3.4.8).

Each of these categories will be discussed below in order to determine whether OGAT can account for these nouns in its current state and, if not, whether any adjustments to the model need to be made.

3.3.4.1 Placenames

There are 11 placenames in the sample. These consist of seven towns/cities, three regions/states and one country.

Towns/cities	Regions/states	Countries
Antwerpen	Bessarabien	Großbritannien
Bautzen	New Hampshire	
Karlovy Vary	Patagonien	
Karlsbad		
Kühlungsborn		
Münster		
Nagasaki		

These 11 placenames are listed without article in the Duden (and therefore without an indication of their gender), since they are not typically accompanied by an article in Standard German: *ich fahre nach Antwerpen* vs. **ich fahre nach dem Antwerpen*. Placenames are, however, used with articles when combined with a modifier, such as a pre-modifying adjective, e.g. *das schöne Antwerpen*, a post-modifying genitive attribute, e.g. *das Antwerpen des Mittelalters*, or a post-modifying prepositional phrase, e.g. *das Antwerpen von morgen* (e.g. Köpcke & Zubin 2005a: 119, Nübling 2015: 318). This is described as “secondary” article usage (e.g. Duden Grammatik 2009: 301, Nübling 2015: 318), since an article is only required by certain contexts (“primary” article usage thus refers to cases in which nouns typically do appear with an article). As is clear from the above examples, *Antwerpen* surfaces with neuter targets in all cases. This is in fact true for all 11 placenames in the sample, e.g. *das atemberaubende Patagonien*, *das Großbritannien von heute*. Supporting evidence in the form of internet search results is presented below.

Placename	No. of Google hits (07/04/17)		
	“der schöne __”	“die schöne __”	“das schöne __”
<i>Antwerpen</i>	0	3 ¹⁹	113
<i>Bautzen</i>	0	0	475
<i>Karlovy Vary</i>	0	0	42
<i>Karlsbad</i>	0	0	152
<i>Kühlungsborn</i>	0	0	142
<i>Münster</i>	0	0	1,980
<i>Nagasaki</i>	0	0	55
<i>Bessarabien</i>	0	0	137
<i>New Hampshire</i>	0	0	100
<i>Patagonien</i>	0	0	316
<i>Großbritannien</i>	0	0	1,810

City/town names

City/town names in fact always surface with neuter targets in German (e.g. Hickey 2000: 630, Menzel 2004: 63, Thomoglou 2004: 34, Hoberg 2004: 106, Köpcke & Zubin 2005a: 119, Weinrich 2007: 328, Duden Grammatik 2009: 301, Engel 2009: 272, Eisenberg 2013: 138, Nübling 2015: 309). This is the case even with city/town names with a formal feature associated with a different gender, e.g. *Halle*_{NEUT}

¹⁹ These three results are from non-German tourist websites and are thus unlikely to have been produced by native German speakers.

(despite $-E \rightarrow *M, *N$), *Jena*_{NEUT} (despite $-A \rightarrow *M, *N$), *Koblenz*_{NEUT} (despite $-ENZ \rightarrow *M, *N$), and compound city/town names with a right-hand element of a different gender, e.g. *Ludwigshafen*_{NEUT} (despite *Hafen*_{MASC}), *Hamburg*_{NEUT} (despite *Burg*_{FEM}), and *Darmstadt*_{NEUT} (despite *Stadt*_{FEM}). Supporting evidence is presented below.

Placename	No. of Google hits (07/04/17) ²⁰		
	“der schöne ___”	“die schöne ___”	“das schöne ___”
<i>Halle</i>	0	0	1,150
<i>Jena</i>	0	0	7,510
<i>Koblenz</i>	0	0	677
<i>Ludwigshafen</i>	0	0	190
<i>Hamburg</i>	0	0	16,800
<i>Darmstadt</i>	0	0	4,520

This suggests that the formal content of city/town names is not analysed by native speakers.

In the case of compound city/town names, there is further evidence indicating that this is true, in that the right-hand elements of these nouns fail many of the tests for headedness (see 3.3.3). On a semantic level, *Ludwigshafen*, for instance, cannot be considered a kind of *Hafen* ‘port’, nor *Hamburg* a kind of *Burg* ‘castle’ (Hoberg 2004: 85, 106f.), thus causing these elements to fail the hyponymy criterion. On a morphological level, not only do such nouns not inherit the gender of their right-hand constituent, but they also do not adopt its plural form: *es gibt zwei Hamburgs* vs. **es gibt zwei Hamburgern*.²¹ Additionally, it is not possible to separate the constituents of such nouns in order to perform morphological operations such as coordination: **morgen fahre ich erst nach Ham- und dann nach Homburg* (Harnisch & Nübling 2004: 193). This indicates that such structures are perceived as simplex.

Consequently, in terms of gender assignment, it might be said that these nouns have purely “semantic gender” or “referential gender” (Dahl 2000: 105f.), since their gender is based entirely on their referent rather than on their formal features.

In the sample, this can be seen with *Kühlungsborn*, which is neuter despite the masculine gender of *Born*, and *Münster*, which is neuter despite its final non-suffixal *-er*, which is associated with masculine (e.g. Steinmetz 1986: 197, Hoberg 2004: 92, Eisenberg 2013: 135).

In its original state, in which all GENDER FEATURES constraints are equally ranked, OGAT is unable to account for the gender of these nouns, as exemplified by the tableau for *Münster* in (38a).

²⁰ Minus any false positives.

²¹ Based on native speaker intuitions and internet search results, e.g. *zwei Hamburgs auf einem Fleck* (Hamburger Abendblatt 07.08.08), *es gibt zwei Hamburgs: das schöne [...], und das andere* (Müller 2011).

(38a)

<i>Münster</i>	GENDER FEATURES		DEFAULT HIERARCHY		
	CITIES/TOWNS→* _M ,* _F	-ER→* _F ,* _N	* _N	* _F	* _M
☞ der Münster	*				*
die Münster	*	*!		*	
das Münster		*!	*		

Indeed, one of the main criticisms of OGAT is that it cannot (always) account for the genders of proper nouns (e.g. Scheibl 2008, see 1.4.5.4).

However, the model can be relatively easily adapted in order to accommodate the purely semantic/referential nature of the gender of city/town names. Specifically, this can be done by the ranking of the constraint CITIES/TOWNS→*_M,*_F above all other GENDER FEATURES constraints. Note that the ranking of this constraint within the “strong” constraints block (see 3.3.3) would not reflect the fact that city/town names are assigned neuter regardless of any formal content, i.e. even if they were to have a “strong” gender-relevant feature associated with another gender. Therefore, the constraint CITIES/TOWNS→*_M,*_F must be ranked above the “strong” constraints block in order to accurately reflect this. With this adaptation, OGAT is able to account for the gender of all seven city/town names in the sample, including *Münster*, as illustrated by (38b).

(38b)

<i>Münster</i>	GENDER FEATURES			DEFAULT HIERARCHY		
	“REFERENTIAL”	“STRONG”	OTHERS	* _N	* _F	* _M
	CITIES/TOWNS→* _M ,* _F		-ER→* _F ,* _N			
der Münster	*!					*
die Münster	*!		*		*	
☞ das Münster			*	*		

The ranking of CITIES/TOWNS→*_M,*_F above the “strong” constraints block will be tested in Chapter 4.

Regions and countries

Names of countries and regions differ from city/town names in German, in that there are some which do not appear with neuter targets, as illustrated by the following examples.

Countries	<i>die Schweiz, die Türkei, die Slowakei, die Mongolei, die Ukraine, die Dominikanische Republik, die Zentralafrikanische Republik, die Vatikanstadt/der Vatikanstaat</i> <i>der Sudan, der Südsudan, der Iran, der Irak, der Jemen, der Niger, der Tschad, der Libanon, der Kosovo</i>
Regions/states	<i>die Lausitz, die Pfalz, die Krim, die Normandie, die Wachau etc.</i>

Such cases are said to be instances of “primary” article usage (e.g. Duden Grammatik 2009: 301, Nübling 2015: 318), since it is usual for an article to accompany these placenames.

Unlike with city/town names, there is evidence of the formal content of some of these country and region names being analysed with regard to gender assignment. For instance, it is likely that the final *-e* of *Ukraine*, the final *-ei* of *Türkei*, *Slowakei* and *Mongolei*, the final *-ie* of *Normandie*, and the head nouns *Staat*_{MASC} and *Stadt*_{FEM} of *der Vatikanstaat* and *die Vatikanstadt* respectively, have influenced the gender assigned to these nouns. This suggests that, unlike city/town names, the gender of region and country names is not purely “semantic gender” or “referential gender”, since it appears to be based on the form of the noun itself.

It is, however, interesting to note that there is evidence of some of the above-listed country names (mainly the masculines) starting to be used with “secondary” neuter targets (Duden Grammatik 2009: 160, Nübling 2015: 314f.), as in the following example:

Um das heutige Irak [...] verstehen zu können (Nordwest Zeitung Online 2017)

Nübling (2015: 314f.) demonstrates that the loss of “primary” non-neuter articles is becoming increasingly common with *Sudan*, *Irak* and *Ukraine*, possibly due to the tendency for placenames to lose their overt article as they gain more official recognition as independent states (e.g. van Langendonck & van der Velde 2009: 1005). Nevertheless, the usage of neuter with these countries is still only a minority occurrence, as demonstrated by the data below.

Placename	No. of Google hits (07/04/17)	
	“der/die heutige ____”	“das heutige ____”
<i>Irak</i>	2,410	722
<i>Sudan</i>	453	135
<i>Ukraine</i>	9,620	8

In terms of OGAT’s ability to account for the genders of the region and country names in the sample, the theory is able to account for all four nouns. Because the gender of regions and countries has been shown not to be entirely “referential” like that of cities/towns, and also because the constraints COUNTRIES→*M,*F and REGIONS→*M,*F cannot be considered “strong”, as not all region and country names are neuter, the constraints COUNTRIES→*M,*F and REGIONS→*M,*F are proposed to be ranked equally alongside other GENDER FEATURES constraints. In this way, OGAT is able to account for the sample data, as shown in (39) to (42).

(39)

	GENDER FEATURES			DEFAULT HIERARCHY		
	“REF.”	“STRONG”	OTHERS			
<i>Großbritannien</i>			COUNTRIES→*M,*F -EN→*F	*N	*F	*M
der <i>Großbritannien</i>			*!			*
die <i>Großbritannien</i>			* *!		*	
☞ das <i>Großbritannien</i>				*		

(40)	GENDER FEATURES			DEFAULT HIERARCHY		
	"REF."	"STRONG"	OTHERS			
<i>Bessarabien</i>			REGIONS→*M,*F -EN→*F	*N	*F	*M
der Bessarabien			*!			*
die Bessarabien			* *!		*	
☞ das Bessarabien				*		

(41)	GENDER FEATURES			DEFAULT HIERARCHY		
	"REFERENTIAL"	"STRONG"	OTHERS			
<i>New Hampshire</i>			REGIONS→*M,*F	*N	*F	*M
der New Hampshire			*!			*
die New Hampshire			*!		*	
☞ das New Hampshire				*		

(42)	GENDER FEATURES			DEFAULT HIERARCHY		
	"REF."	"STRONG"	OTHERS			
<i>Patagonien</i>			REGIONS→*M,*F -EN→*F	*N	*F	*M
der Patagonien			*!			*
die Patagonien			* *!		*	
☞ das Patagonien				*		

Additionally, should it be the case that all country names do come to be used with neuter targets, as e.g. Nübling (2015) indicates, then this can be accounted for in OGAT by the ranking of the constraint COUNTRIES→*M,*F within the "referential" block.

3.3.4.2 Personal names

The sample also contains 11 names used for people. These are either first-names, surnames or bynames of well-known figures.

<i>Iskariot</i>	<i>Naftali</i>	<i>Sophia</i>
<i>Klimt</i>	<i>Rahel</i>	<i>Tintoretto</i>
<i>Lärtes</i>	<i>Schiele</i>	<i>Walter</i>
<i>Nadja</i>	<i>Sigrun</i>	

As with the placenames in the sample, all of these names are listed without an article in the Duden because they typically appear without an article in Standard German: *Nadja sprach mit Sophia* (e.g. Duden Grammatik 2009: 299, Nübling 2015: 306). However, names often appear with an article when combined with modifiers, e.g. *die kluge Sigrun*, *der Klaus vom Theater* (Duden Grammatik 2009: 301, Nübling 2015: 318). Occasionally, articles are also used with names to express either distance or derogation, e.g. *Goethe oder nur der Schiller?* (Duden Grammatik 2009: 301). Furthermore, in Northwestern and Southern Germany, names are frequently accompanied by an article in spoken language, e.g. *ich bin die Sophia* (Elspaß & Möller 2003a, 2003b; Duden Grammatik 2009: 301).

In such contexts, names referring to men are consistently used with masculine targets, and names referring to women are consistently used with feminine targets (e.g. Duden Grammatik 2009: 154).²² Crucially, this is true even in cases where a name has a formal feature associated with a different grammatical gender, e.g. a man named *Uwe* would be referred to as *der Uwe* (despite $-E \rightarrow *M, *N$) and a woman with the surname *Bauer* might be referred to as *die Bauer* (despite $-ER \rightarrow *F, *N$) (Nübling 2015: 307). This is also the case with compound names, e.g. *die Almut* (despite Mut_{MASC}) or *die Zimmermann*, when referring to a woman (despite $Mann_{MASC}$) (Nübling 2015: 307). Additionally, this is true in cases where the same name is used to refer to both men and women, e.g. gender-neutral first-names, such as *Kai/Kay*, *Uli*, *Tomke* and *Kirsten* (Duden Online Wörterbuch 2018), e.g. *der kleine Uli* vs. *die kleine Uli* (Köpcke & Zubin 2005a: 120), and all surnames: *da kommt der Schmidt* (a man) vs. *da kommt die Schmidt* (a woman).

Therefore, as was established for cities/towns in 3.3.4.1, the formal content of personal names is not analysed by native speakers with regard to gender assignment. Personal names can thus also be said to have purely “semantic gender” or “referential gender” (Dahl 2000: 105f.), since their gender is based entirely on their referent rather than on their form (e.g. Plank 1986: 58, Köpcke & Zubin 2005a: 120).

Accordingly, it is proposed that the constraints $NAME:MAN \rightarrow *F, *N$ and $NAME:WOMAN \rightarrow *M, *N$ are ranked alongside $CITIES/TOWNS \rightarrow *M, *F$ in the highly-ranked “referential” block.²³ This reflects the fact that personal names will be assigned gender according to their referent irrespective of the formal content of the noun (including if the noun contained a feature associated with a “strong” constraint, see 3.3.3).

In this way, OGAT is able to account for the data in the sample. It can, for instance, determine the feminine gender of *Nadja* if this name is used to refer to a woman (as shown in (43)) and the masculine gender of *Klimt*, if this is used to refer to a man (as shown in (44)). It can also determine the feminine gender of *Klimt*, should the name be used to refer to a woman (as shown in (45)).

²² Women’s names are also found with neuter targets in unmarked contexts in some German dialects, e.g. *das Anna* (e.g. Christen 1998, Nübling et al. 2013). Additionally, there are cases in Standard German where women’s names are found with neuter targets, e.g. *das Merkel*. However, these cases are highly marked and used to express derogation (e.g. Nübling 2014). Interestingly, there is also some usage of neuter with names within the non-binary community (e.g. Lohr 2016, 2017).

²³ It also seems possible that the “strong” semantic constraints identified in 3.3.3.3, e.g. $CHEESES \rightarrow *F, *N$; $LANGUAGES \rightarrow *M, *F$ and $METALS \rightarrow *M, *F$, are also ranked in the “referential” block. This will be investigated in Chapter 4. Note that the relationship between constraints within the “referential” block is irrelevant, since it will never be the case that more than one of these constraints applies to any given noun.

(43)		GENDER FEATURES			DEFAULT HIERARCHY		
		"REFERENTIAL"	"STRONG"	OTHERS	*N	*F	*M
	<i>Nadja</i>	NAME:WOMAN→*M,*N		-A→*M,*N	*N	*F	*M
	der Nadja	*!					*
	☞ die Nadja					*	
	das Nadja	*!			*		

(44)		GENDER FEATURES			DEFAULT HIERARCHY		
		"REFERENTIAL"	"STRONG"	OTHERS	*N	*F	*M
	<i>Klimt</i>	NAME:MAN→*F,*N		CCVCC -N(c)(c) →*F,*N →*F,*N	*N	*F	*M
	☞ der Klimt						*
	die Klimt	*!		* *		*	
	das Klimt	*!		* *	*		

(45)		GENDER FEATURES			DEFAULT HIERARCHY		
		"REFERENTIAL"	"STRONG"	OTHERS	*N	*F	*M
	<i>Klimt</i>	NAME:WOMAN→*M,*N		CCVCC -N(c)(c) →*F,*N →*F,*N	*N	*F	*M
	der Klimt	*!					*
	☞ die Klimt			* *		*	
	das Klimt	*!		* *	*		

There is, however, one case relating to personal names for which the modified version of OGAT currently cannot account. This is the fact that personal names which are suffixed with *-chen* (or *-lein* or *-le*) are always neuter, regardless of the societal gender of the referent, e.g. *das Lottchen*, *das Fritzchen*, *das Ännchen* (e.g. Heringer 1995: 214, Harnisch & Nübling 2004: 1906f., Duden Grammatik 2009: 154). This therefore appears to be an instance of the formal content of a name being analysed and determining gender assignment, i.e. a case in which the gender of a name is not purely "referential".

Consequently, our model of OGAT needs further adaptation, since, in its current state, *Fritzchen*, for instance, would be assigned masculine, as shown in (46a). In order to account for the data, OGAT requires the constraint *-CHEN→*M,*F* to be ranked above the "referential gender" constraints, as shown in (46b).

(46a)		GENDER FEATURES			DEFAULT HIERARCHY		
		"REFERENTIAL"	"STRONG"	OTHERS	*N	*F	*M
	<i>Fritzchen</i>	NAME:MAN→*F,*N	-CHEN→*M,*F		*N	*F	*M
	☞ der Fritzchen		*				*
	die Fritzchen	*!	*			*	
	das Fritzchen	*!			*		

(46b)

	GENDER FEATURES				DEFAULT HIERARCHY		
	DIMINUTIVE	“REFERENTIAL”	“STRONG”	OTHERS	*N	*F	*M
<i>Fritzchen</i>	-CHEN→*M,*F	NAME:MAN→*F,*N					
der Fritzchen	*!						*
die Fritzchen	*!	*				*	
☞ das Fritzchen		*			*		

It is, however, important to make clear that final *-chen* is only a gender-determining feature of a name when it is analysed as a morpheme, which might be illustrated by the insertion of a morpheme boundary, as in: +CHEN→*M,*F, or simply by specifying this in the constraint definition. For instance, the surname *Schmidtchen* as in *da kommt der/die Schmidtchen* (Herr/Frau Schmidtchen) is simplex, i.e. not analysed as containing a diminutive suffix, and therefore not subject to the constraint -CHEN→*M,*F. It thus receives masculine/feminine gender due to the constraints in the “referential” block. Conversely, the surname *Schmidt* when suffixed with *-chen* in an utterance such as *das kleine Schmidtchen* surfaces as neuter, because its final *-chen* is analysable a diminutive suffix, and is therefore subject to the constraint -CHEN→*M,*F in the same way as *Fritzchen* in (48b).

The proposals made in this section will be tested in Chapter 4.

3.3.4.3 Plural terms

In the sample, there are four nouns listed in the Duden in their plural form without any indication of their gender.

Noun	Singular form
<i>Gerichtskosten</i> ‘legal expenses’	–
<i>Frutti</i> ‘fruit’	–
<i>Korpuskularstrahlen</i> ‘corpuscular radiation’	<i>der Strahl</i>
<i>Geisa</i> ‘geisa’ (architectural feature of ancient Greek and Roman buildings)	<i>das Geison</i>

Kosten ‘cost’ is a *plurale tantum*, i.e. a noun with no singular form. Its gender cannot therefore be determined, since there is no gender marking on plural targets in German (see 1.1).²⁴ Similarly, there is no corresponding singular form for *Frutti* listed in the Duden. The Italian singular form of the loan, *frutto*, does not appear to be in usage in German. Consequently, OGAT cannot be tested on these nouns.

²⁴ Note that *die Kost* ‘food’ cannot, at least synchronically, be considered the singular form of *Kosten*.

The remaining two nouns do, however, have corresponding singular forms, which are given a separate lexical entry in the Duden. *Das Geison* is the singular of *Geisa*, and *der Strahl* is the singular of *Strahlen*. OGAT is able to assign the correct gender in both cases.

(47)

	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS				
<i>Geison</i>	-ON UNSTRESS. → *M, *F		*N	*F	*M
der <i>Geison</i>	*!				*
die <i>Geison</i>	*!			*	
☞ <i>das Geison</i>			*		

(48)

	GENDER FEATURES					DEFAULT HIERARCHY		
	OTHERS							
<i>Strahl</i>	[[C- MONO → *F, *N	-L MONO → *F	V _{LONG} (C) MONO → *F	CCCVC → *F, *N	(DEVERBAL → *F, *N) ²⁵	*N	*F	*M
☞ <i>der Strahl</i>								*
die <i>Strahl</i>	*	*	*	*!	(*)		*	
das <i>Strahl</i>	*			*!	(*)	*		

3.3.4.4 Regional terms

There are nine nouns in the sample which are listed in the Duden as regional terms.

Noun	Area
<i>die Angelobung</i> 'inauguration'	Austria
<i>der Außerstreitrichter</i> 'non-adversarial proceedings judge'	Austria
<i>das Außerstreitverfahren</i> 'non-adversarial proceedings'	Austria
<i>der Detailhandel</i> 'retail trade'	Switzerland
<i>der Farre</i> 'young bull'	Unspecified
<i>das Flett</i> 'living/kitchen area in old Lower Saxon farmhouse'	Lower Saxony
<i>der Rahm</i> 'cream'	Austria, Southern Germany, Western Central Germany, Switzerland
<i>der Sigrüst</i> 'parish clerk'	Switzerland
<i>die Tramway</i> 'tram'	Eastern Austria, Southern Germany

In its current state, OGAT is able to assign the correct gender to eight of these nouns, as exemplified by *Sigrüst* in (49).

(49)

	GENDER FEATURES		DEFAULT HIERARCHY		
	"STRONG" CONSTRAINTS	OTHERS			
<i>Sigrüst</i> 'parish clerk'	-IST → *F, *N	GEN. PERSON → *F, *N	*N	*F	*M
☞ <i>der Sigrüst</i>					*
die <i>Sigrüst</i>	*!	*		*	
das <i>Sigrüst</i>	*!	*	*		

²⁵ See Section 3.2.2.1.

OGAT cannot, however, currently account for the neuter gender of *Flett*. In order to resolve this, the ranking $-ETT \rightarrow *M, *F \gg CCVC \text{ MONO} \rightarrow *F, *N$ is necessary within the GENDER FEATURES block, as demonstrated in (50). This ranking does not affect any other nouns in the sample, but is consistent with other nouns in the German lexicon such as *das Brett* and *das Frett*. This ranking will be tested in Chapter 4.

(50a)	<i>Flett</i>	GENDER FEATURES		DEFAULT HIERARCHY		
		OTHERS		*N	*F	*M
		$-ETT \rightarrow *M, *F$	$CCVC \text{ MONO} \rightarrow *F, *N$			
<i>der Flett</i>	*				*	
<i>die Flett</i>	*	*!		*		
<i>das Flett</i>		*	*!			

(50b)	<i>Flett</i>	GENDER FEATURES		DEFAULT HIERARCHY		
		OTHERS		*N	*F	*M
		$-ETT \rightarrow *M, *F$	$CCVC \text{ MONO} \rightarrow *F, *N$			
<i>der Flett</i>	*!				*	
<i>die Flett</i>	*!	*		*		
<i>das Flett</i>		*	*			

3.3.4.5 Archaic nouns

The sample contains three nouns which are described in the Duden as archaic.

<i>die Metropolis</i>	archaic term for <i>Metropole</i> 'metropolis'
<i>der Tip</i>	archaic term for <i>Tipp</i> 'hint'
<i>die Elektrische</i>	archaic term for <i>Straßenbahn</i> 'tram'

OGAT is able to account for the gender of *Metropolis*, as shown in (51).

(51)	<i>Metropolis</i>	GENDER FEATURES		DEFAULT HIERARCHY		
		OTHERS		*N	*F	*M
		$-IS \rightarrow *M, *N$				
<i>der Metropolis</i>		*!			*	
<i>die Metropolis</i>				*		
<i>das Metropolis</i>		*!	*			

OGAT is also able to account for *Tip*, which is simply an archaic spelling of *Tipp*.

(52)	<i>Tip(p)</i>	GENDER FEATURES		DEFAULT HIERARCHY		
		OTHERS		*N	*F	*M
		$T- \text{MONO} \rightarrow *F$	$CVC \text{ MONO} \rightarrow *F$			
<i>der Tip</i>					*	
<i>die Tip</i>	*	*!		*		
<i>das Tip</i>			*!			

Finally, *Elektrische* is an adjectival noun, and is thus handled in 3.3.4.7.

3.3.4.6 Homonyms

When collecting the sample, five nouns were encountered whose lemmata are linked in the Duden since they are classified as homonyms.

<i>der Aviso</i> (1) 'small warship' <i>das Aviso</i> (2) 'notification'	<i>der Holländer</i> (1) 'Dutchman' <i>der Holländer</i> (2) 'Dutch cheese' <i>der Holländer</i> (3) '4-wheeled child's bike (of Dutch origin)'
---	---

In 3.3.2.1.4, it was established that *der Aviso* and *das Aviso* are indeed homonyms (rather than polysemes) due to their divergent etymologies, lack of semantic relatedness and different morphosyntactic properties, e.g. gender. Using the same criteria (see 3.3.2.1.3), the three entries for *der Holländer* might, however, be argued to be cases of polysemy rather than homonymy, despite their classification as homonyms in the Duden. This is due to the relatedness between the senses (i.e. all of Dutch origin) and shared gender.

In terms of gender assignment, the categorisation of *Holländer* is, however, unimportant, since OGAT is able to assign the correct gender to all three instances of the noun whether they are treated as homonyms or as a polyseme (see 3.3.2.1.6), as shown in (53).

(53a)

Holländer 'Dutchman'	GENDER FEATURES		DEFAULT HIERARCHY		
	"STRONG" CONSTRAINTS	OTHERS	*N	*F	*M
	-ER SUFF. → *F, *N	GEN.PERSON → *F, *N			
<i>der Holländer</i>					*
<i>die Holländer</i>	*!	*		*	
<i>das Holländer</i>	*!	*	*		

(53b)

Holländer 'cheese'	GENDER FEATURES		DEFAULT HIERARCHY		
	"STRONG" CONSTRAINTS	OTHERS	*N	*F	*M
	-ER SUFF. → *F, *N	CHEESES → *F, *N ²⁶			
<i>der Holländer</i>					*
<i>die Holländer</i>	*	*!		*	
<i>das Holländer</i>	*	*!	*		

²⁶ As mentioned in 3.3.4.2, it is possible that this constraint (along with other "strong" semantic constraints identified in 3.3.3.3) is instead ranked in the "referential gender" block. This will be tested in Chapter 4.

(53c)

	GENDER FEATURES		DEFAULT HIERARCHY		
	"STRONG" CONSTRAINTS	OTHERS			
Holländer 'child's bike'	-ER SUFF. → *F, *N		*N	*F	*M
☞ <i>der Holländer</i>					*
<i>die Holländer</i>	*!			*	
<i>das Holländer</i>	*!		*		

For *der Aviso* and *das Aviso*, however, OGAT is able to account for the gender of *der Aviso* only, as shown below.

(54)

	GENDER FEATURES			DEFAULT HIERARCHY		
	"STRONG"	OTHERS	SEM. EQUIV.			
Aviso 'warship'		-O → *F		*N	*F	*M
☞ <i>der Aviso</i>						*
<i>die Aviso</i>		*!			*	
<i>das Aviso</i>				*!		

(55)

	GENDER FEATURES			DEFAULT HIERARCHY		
	"STRONG"	OTHERS	SEM. EQUIV.			
Aviso 'notification'		-O → *F	SEM. EQUIV (DER/DAS AVIS) → *F/ (DER HINWEIS) → *F, *N ²⁷	*N	*F	*M
☞ <i>der Aviso</i>						*
<i>die Aviso</i>		*!	*		*	
<i>das Aviso</i>			(*!)	*(!)		

It could be argued that for *das Aviso*, a kind of blocking effect is in force due to homonymy avoidance, which might be represented by a constraint such as EXISTING HOMONYM → *M. This would enable OGAT to correctly assign neuter. However, homonymy blocking as an explanation for such phenomena has been heavily criticised (see e.g. Clark 1993, Plag 2003: 79f., Enger 2007). Furthermore, such a constraint would require more thorough testing on a significantly wider range of homonyms.

3.3.4.7 Adjectival nouns

There are four adjectival nouns which were included in the sample as outlined in 2.2.

<i>der/die Kranke</i> 'ill man/woman'	<	<i>krank</i> 'ill, sick'
<i>der/die Renitente</i> 'unruly man/woman'	<	<i>renitent</i> 'unruly'
<i>das Sechsfache</i> 'sixfold'	<	<i>sechsfach</i> 'sixfold'
<i>die Elektrische</i> '(electric) tram'	<	<i>elektrisch</i> 'electric'

²⁷ As discussed in 3.3.1.2, both *der Hinweis* (the most frequent equivalent) and *der/das Avis* (a cognate) are plausible semantic equivalents to *das Aviso*. It is unclear which is predominant, but in either case the outcome is identical.

Typically, adjectival nouns in German can be used with all three genders. In a similar manner to cities/towns and personal names (see 3.3.4.1 and 3.3.4.2), the formal content of these elements is not analysed, and their gender is based on their referent. When used to refer to a man, adjectival nouns surface as masculine, when used to refer to a woman, they surface as feminine, and when used to refer to an abstract concept, an inanimate or a young person/animal, they surface as neuter (e.g. Duden Grammatik 2009: 349). In the sample, this can be seen with *der/die Kranke* ‘ill man/woman’, *der/die Renitente* ‘unruly man/woman’ and *das Sechsfache* ‘sixfold’.

The gender of such constructions may not usually be expected to be accounted for in models of gender assignment, given that their status as nouns is unclear (e.g. Schwarze 2008: 155). Technically, however, OGAT could account for the genders of these elements by means of constraints in the “referential gender” block (see 3.3.4.1 and 3.3.4.2).

Additionally, there are some fixed elliptical constructions in which adjectival nouns appear, where the gender of the adjectival noun is determined by the elided head, e.g. *ein Helles (Bier)*. In the sample, this can be seen with *die Elektrische* ‘tram’, whose elided head is *die (Straßen)bahn*. Again, OGAT would be able to account for the gender of this element through reference to the head (see 3.3.3).

3.3.4.8 Multiple-gender nouns

Finally, there are eight multiple-gender nouns (MGNs) in the sample. As stated in 1.1, these are nouns which can be used with more than one gender without any effect on the noun’s meaning (e.g. Talanga 1987: 14, Schulte-Beckhausen 2002: 79f., Onysko 2007: 174), and which constitute just 1.3% of nouns in the Duden corpus (Duden Sprachwissen 2017) and 1.4% of nouns in the sample (see 2.2).

<i>der/das Renforcé</i> ‘renforcé [fabric]’	<i>der/das Kulm</i> ‘rounded summit’	<i>das/der Ar</i> ‘are [unit]’
<i>der/das Juchten</i> ‘Russia leather; fragrance’	<i>der/das New Look</i> ‘new look’	<i>das/die Korpuskel</i> ‘corpuscle’
<i>der/das Garnknäuel</i> ‘ball of yarn’	<i>der/das Avis</i> ‘notification’	

All eight MGNs in the sample are listed as belonging to two genders (as opposed to three), like the vast majority (99.98%) of MGNs in the Duden (Duden Sprachwissen 2017). Seven of these nouns are listed as being masculine or neuter, like 70% of MGNs in the Duden, and one is listed as feminine or neuter, like 12% of MGNs in the Duden (Duden Sprachwissen 2017).

A number of observations have been made with regard to the variation in gender of MGNs in German. Firstly, it is stated that gender may vary according to region (e.g. Hickey 2000: 625, Di Meola & Puato 2017: 33). In the sample, this can be seen with *Ar*, which is listed as neuter in Austria but both neuter

and masculine in Germany. Callies et al. (2012) hypothesise that such variation in gender might be reinforced by varying institutionalised preferences in different regions.

Secondly, it has been claimed that the gender of a MGN may vary according to register (Duden Grammatik 2009, Di Meola & Puato 2017: 33f.) In the sample, this can be seen with *das/die Korpuskel*, which is said to be neuter in standard contexts but feminine in technical language (Duden Online-Wörterbuch 2018d).

Thirdly, it is noted that many MGNs are nouns which are recognisable as (relatively recent) loanwords (e.g. Schulte-Beckhausen 2002, Hoberg 2004: 79, Callies et al. 2012). In the sample, this is true for *Avis*, *Renforcé* and *New Look*. It is generally thought that this is due to the loans not yet being fully integrated in the host language (e.g. Poplack et al. 1982, Poplack & Sankoff 1984: 124, Hoberg 2004: 79), and/or due to the loans being subject to conflicting gender assignment principles (including having competing semantic equivalents) (e.g. Di Meola & Puato 2017: 34, Franco et al. 2018), and/or due to variation in the level of knowledge of the donor language among native German speakers (e.g. Fischer 2005). It has also been claimed that there might be regional variation among loans due to language contact (e.g. with French) (e.g. van der Elst 1983: 1204ff., Schulte-Beckhausen 2002: 38).

Finally, it has been observed that nouns with low frequencies are more likely to be MGNs than highly frequent nouns (e.g. Talanga 1987: 133, Di Meola & Puato 2017: 34). This is in line with the finding that, according to the Leipzig Corpora Collection and DWDS, all eight MGNs in the sample belong to among the lowest frequency classes, with less than 50 tokens in total found for *Renforcé*, *Korpuskel*, *Juchten* and *Kulm*. Related to this is the claim that MGNs are more likely to occur among nouns which are used without an article or in other contexts where the gender marking is not distinctive (Di Meola & Puato 2017: 34). This is particularly relevant for the seven masculine/neuter MGNs in the sample due to the considerable syncretism in the paradigms of masculine and neuter targets.

It is now important to investigate whether OGAT in its current state has the ability to account for the genders of the MGNs in the sample.

Firstly, the loanwords *Avis*, *New Look* and *Renforcé* will be examined. For *Avis* and *New Look*, the variation in gender might simply be accounted for by their competing semantic equivalents (as stated by e.g. Di Meola & Puato (2017: 34), see also 3.3.1.2). For *Avis*, this might be the competition between the most frequent equivalent *der Hinweis* and the cognate equivalent *das Aviso*, and for *New Look*, this might be the competition between the most frequent equivalent *der Stil* and the second most frequent equivalent *das Aussehen* (see 3.3.1.2), as shown in (56) and (57).

(56a)

Avis	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS	SEM. EQUIV.	*N	*F	*M
		SEM.EQUIV. → *F, *N (DER HINWEIS)			
<i>der Avis</i>					*
<i>die Avis</i>		*!		*	
<i>das Avis</i>		*!	*		

(56b)

Avis	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS	SEM. EQUIV.	*N	*F	*M
		SEM.EQUIV. → *M, *F (DAS AVISO)			
<i>der Avis</i>		*!			*
<i>die Avis</i>		*!		*	
<i>das Avis</i>			*		

(57a)

New Look	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS	SEM. EQUIV.	*N	*F	*M
	CVC → *F	SEM.EQUIV. → *F, *N (DER STIL)			
<i>der New Look</i>					*
<i>die New Look</i>	*!	*		*	
<i>das New Look</i>		*!	*		

(57b)

New Look	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS	SEM. EQUIV.	*N	*F	*M
	CVC → *F	SEM.EQUIV. → *M, *F (DAS AUSSEHEN)			
<i>der New Look</i>		*!			*
<i>die New Look</i>	*!	*		*	
<i>das New Look</i>			*		

It is interesting to observe that Wahrig (1997) lists *Avis* not as a MGN but as neuter, indicating the dominance of the cognate equivalent over the most frequent equivalent (in line with Geerts (1996), who states that cognates are likely to be more influential than non-cognate equivalents). Similarly, both Wahrig (1997) and the DWDS, as well as around half of the 32 dictionaries examined in Schulte-Beckhausen (2002), list *New Look* as masculine rather than a MGN. This is in line with the fact that the most frequent equivalent is *der Stil* (see 3.3.1.2).

For *Renforcé*, the variation in gender might be accounted for by variation in the knowledge of the donor language (as stated by e.g. Fischer 2005). This is because *Renforcé* is the product of the nominal conversion of the French past participle *renforcé* (from *renforcer*). As stated in 1.3.1.3, nominal conversions in German are typically neuter. Therefore, speakers with knowledge of French might be more likely to be aware of its status as a converted past participle and apply the constraint

CONVERSION→*M,*F, as in (58b). Those without such knowledge are more likely to produce the analysis shown in (58a).²⁸

(58a)

Renforcé	GENDER FEATURES		DEFAULT HIERARCHY		
	"STRONG"	OTHERS	*N	*F	*M
		-EE/-ÉE/-É→*F FABRIC→*F,*N			
<i>der Renforcé</i>					*
<i>die Renforcé</i>	*	*!		*	
<i>das Renforcé</i>		*!	*		

(58b)

Renforcé	GENDER FEATURES		DEFAULT HIERARCHY		
	"STRONG"	OTHERS	*N	*F	*M
	CONVERSION→*M,*F	-EE/-ÉE/-É→*F FABRIC→*F,*N			
<i>der Renforcé</i>	*!				*
<i>die Renforcé</i>	*!	* *		*	
<i>das Renforcé</i>		*	*		

For *der/das Juchten*, the variation in gender might be accounted for by the variation in meaning of the noun. According to the Duden, the noun is polysemous, with the two senses: (1) type of leather, and (2) fragrance (similar to that of the leather). According to the first sense (which is arguably also its primary, prototypical sense, see 3.3.2.1), the noun is subject to the constraint LEATHER TYPES→*M,*F (e.g. Schulte-Beckhausen 2002: 73, Chan 2005: 47), which causes OGAT to predict neuter for *Juchten*, as in (59a). According to the second sense, however, this constraint does not apply, causing OGAT to predict masculine, as in (59b). Therefore, it is possible that the gender variation could be accounted for by the speakers' perception of *der* and *das Juchten* as homonyms, rather than as a polyseme (where only the primary, prototypical sense is relevant to gender assignment, see 3.3.2.1). Interestingly, however, Wahrig (1997) lists *Juchten* not as a MGN but as neuter, suggesting that (59a) may be the more common analysis,²⁹ which is compatible with the noun's perception as a polyseme.

(59a)

Juchten	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS		*N	*F	*M
	-EN→*F	LEATHER TYPES→*M,*F			
<i>der Juchten</i>		*!			*
<i>die Juchten</i>	*	*!		*	
<i>das Juchten</i>			*		

²⁸ Some speakers assigning masculine to *Renforcé* who have knowledge of French might also be influenced by the fact that nouns ending in *-é* in French are typically masculine (e.g. Rosenthal 2009: 58).

²⁹ Note that the frequency of the occurrence of the MGNs in the sample with each of their possible genders could not be reliably ascertained using the corpora available, primarily due to the scarcity of data (given that these are mostly highly infrequent nouns) and/or the abundance of false positives and repeated results.

(59b)

Juchten	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS		*N	*F	*M
	-EN→*F				
<i>der Juchten</i>					*
<i>die Juchten</i>		*!		*	
<i>das Juchten</i>			*!		

For *Ar*, *Korpuskel* and *Kulm*, OGAT is able to predict only one of the two possible genders of these nouns. However, in each case, there is reason to believe that this is the “primary” gender of the two possibilities. For *Ar*, the Duden entry states ‘das, österreichisch nur so, auch: der [das, in Austrian German only so, also: der]’ and Wahrig (1997) states ‘das, selten der [das, rarely der]’, thus suggesting that neuter may be the more usual gender, which is indeed the gender that OGAT predicts. Interestingly, the second most optimal gender for *Ar* according to OGAT would be masculine.

(60)

Ar	GENDER FEATURES				DEFAULT HIERARCHY		
	OTHERS				*N	*F	*M
	SCI.UNITS→*M,*F	VC→*M,*F	V _{LONG} (C)MONO→*F	∅V _{LONG} MONO→*F			
<i>der Ar</i>	*	*!				*	
<i>die Ar</i>	*	*	*	*!		*	
<i>das Ar</i>					*		

For *Korpuskel*, the Duden entry states ‘das (fachsprachlich auch: die) [das (in technical language also: die)]’, and the DWDS lists only neuter as a possibility, thus indicating that neuter is the more standard gender, which, again, is the gender that OGAT predicts.

(61)

Korpuskel	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS		*N	*F	*M
	SCI. UNITS→*M,*F				
<i>der Korpuskel</i>		*!			*
<i>die Korpuskel</i>		*!		*	
<i>das Korpuskel</i>			*		

For *Kulm*, the Duden allows both masculine and neuter, yet Wahrig (1997) lists only the masculine variant, which is the gender that OGAT predicts. Interestingly, as with *Ar*, the other possible gender (in this case neuter) is considered the second most optimal gender according to OGAT.

(62)

Kulm	GENDER FEATURES			DEFAULT HIERARCHY		
	OTHERS			*N	*F	*M
	CVCC→*F	-[!]C MONO→*F	-N(C)(C) MONO→*F,*N			
<i>der Kulm</i>						*
<i>die Kulm</i>	*	*	*!		*	
<i>das Kulm</i>			*!	*		

Finally, for *Garnknäuel*, OGAT is able to predict only masculine. There is no apparent synchronic motivation for the variability in gender of this noun; there are no known, valid GENDER FEATURES constraints which are relevant to the noun, so OGAT assigns masculine by default (note that default genders are investigated further in 3.4.3). There is no reason to believe that this is the “primary” gender of the noun, since both genders are listed in various dictionaries, and corpus data reveal usage of the noun with both genders.

(63)

	GENDER FEATURES	DEFAULT HIERARCHY		
	OTHERS	*N	*F	*M
Garnknäuel				
<i>☞ der Garnknäuel</i>				*
<i>die Garnknäuel</i>			*!	
<i>das Garnknäuel</i>		*!		

Additionally, it is important to note that the oft-cited methods proposed for accounting for variation in OT, such as floating constraints (e.g. Reynolds 1994, Nagy & Reynolds 1997), are largely incompatible with OGAT, and in most cases would over-generate variation. These therefore cannot be used to account for the multiple genders of *Garnknäuel*.

In sum, it has been demonstrated that for all eight MGNs in the sample, OGAT is able to account for at least one of the possible genders of the nouns. In half of the cases, OGAT is able to account for the variation in gender through the optionality of certain GENDER FEATURES constraints. For the remaining nouns, OGAT is able to predict only one of their possible genders, although in three cases there is some indication that this gender can be considered the “primary” gender. It would be interesting to see whether, in time, the gender that OGAT predicts for these nouns matches the gender that the nouns come to be used with exclusively, assuming that this will indeed occur.³⁰

3.4 Remaining issues

In the above sections it has been established that OGAT in its original state is able to account for the genders of many nouns in the sample (3.1), and that it is able to account for an even greater proportion of nouns if various adaptations to the theory are made (3.2–3.3). In this section, a number of remaining issues are considered, in order to establish whether further modifications to OGAT are needed if it is to account for the genders of all nouns in the sample.

³⁰ There is ample evidence to suggest that this is the common trajectory for MGNs (e.g. *der/das/die Dschungel* > *der Dschungel* (Duden Grammatik 2009: 167)). Nevertheless, there are also a few cases of MGNs being used with multiple genders over extended periods (e.g. MHG *das/der zepter/cepter* > NHG *das/der Zepter* (Duden Grammatik 2009: 169)).

The first issue involves the proposal of a small number of new GENDER FEATURES constraints which are required by OGAT to account for the genders of various nouns in the sample (3.4.1). The second issue involves the ranking of a small number of individual GENDER FEATURES constraints in order to improve OGAT's ability to account for the sample data (3.4.2). The third issue involves the investigation of possible alternative rankings within the default markedness hierarchy, as discussed in 1.4.5.3 (3.4.3).

3.4.1 New constraints

As noted by McCarthy (2008: 169), when performing an analysis in OT, it is occasionally necessary to posit new constraints, since the existing constraint set may prove inadequate. Indeed, there are 11 nouns in the sample for which OGAT in its current state cannot account, but could be easily accounted for with the addition of a small number of GENDER FEATURES constraints.³¹

<i>die Anstalt</i>	<i>das Arrival</i>	<i>der Asphalt</i>
<i>das Deleatur</i>	<i>das Eclair</i>	<i>das Filet</i>
<i>die Garonne</i>	<i>das Informel</i>	<i>der Islam</i>
<i>das Partizip</i>	<i>das Rondeau</i>	

The proposed new constraints are listed below.

Constraint	Description	Nouns in sample
-AIR [ɛ:p̥]→*M,*F	Nouns ending in <i>-air</i> [ɛ:p̥] cannot be masculine or feminine, e.g. <i>das Flair</i> , <i>das Pleinair</i>	<i>Eclair</i>
-AL ENG.→*M,*F	Nouns ending in <i>-al</i> , where the <i>-al</i> is unstressed (usually realised as either [ə] or [ɪ]) cannot be masculine or feminine, e.g. <i>das Musical</i> , <i>das Revival</i>	<i>Arrival</i>
-ALT→*N	Nouns ending in <i>-alt</i> cannot be neuter, e.g. <i>der Inhalt</i> , <i>die Gewalt</i>	<i>Asphalt</i> , <i>Anstalt</i>
-EAU→*M,*F	Nouns ending in <i>-eau</i> [ˈo:] cannot be masculine or feminine, e.g. <i>das Niveau</i> , <i>das Tableau</i>	<i>Rondeau</i>
-EL(L) [ˈɛl]→*M,*F	Nouns ending in stressed <i>-el(l)</i> [ˈɛl] cannot be masculine or feminine, e.g. <i>das Bordell</i>	<i>Informel</i>
-ET→*M,*F	Nouns ending in stressed <i>-et</i> [ˈe:] cannot be masculine or feminine, e.g. <i>das Bidet</i> , <i>das Buffet</i>	<i>Filet</i>
-IP [i:p]→*M,*F	Nouns ending in <i>-ip</i> [i:p] cannot be masculine or feminine, e.g. <i>das Prinzip</i>	<i>Partizip</i>
-NNE [n(ə)]→*M,*N	Nouns ending in <i>-nne</i> [V _{STRESSED:SHORT} n(ə)] cannot be masculine or neuter, e.g. <i>die Chaconne</i> , <i>die Egyptienne</i>	<i>Garonne</i>
RELIGIONS→*F,*N	Nouns denoting religions/spiritual traditions cannot be feminine or neuter, e.g. <i>der Hinduismus</i> , <i>der Voodoo</i>	<i>Islam</i>
SYMBOLS→*M,*F	Nouns denoting typographical symbols and punctuation marks cannot be masculine or feminine e.g. <i>das Komma</i> , <i>das Hashtag</i> , <i>das Hatschek</i>	<i>Deleatur</i>

³¹ Note that numerous other new constraints were also tested which would have enabled OGAT to account for the genders of the nouns listed in the section on exceptions (3.5.2). However, these failed to meet the criteria and are thus not discussed here.

McCarthy (2008: 169) states that where new constraints are necessary, these must be properly defined and their motivation explained, as done above. Additionally, in line with the reasoning outlined in 1.4.5.3, it is necessary to test new constraints using Enger’s (2009) criteria in order to determine whether there is independent evidence to support the proposals. This will be done using the methods established in Chapter 2.

Criterion 1: Share of candidate nouns

The table below shows the proportion of nouns whose gender can be accounted for by the proposed constraints.

Constraint	Candidate nouns in sample	Exceptions	Exceptions tolerated	Valid according to TP?	% Coverage ↓
-IP [i:p]→*M,*F	2	0	2.0 ³²	✓	100.0
-ALT→*N	24	3	9.9	✓	87.5
-AL ENG.→*M,*F	15	2	7.0	✓	86.7
-ET→*M,*F	27	4	10.9	✓	85.2
-NNE [n(ə)]→*M,*N	43	8	18.1	✓	84.3
-EAU→*M,*F	12	2	6.0	✓	83.3
-AIR [ɛ:ɐ]→*M,*F	9	2	4.9	✓	77.8
RELIGIONS→*F,*N	74	17	24.6	✓	77.0
-EL(L) [ʼɛl]→*M,*F	41	11	15.1	✓	73.2
SYMBOLS→*M,*F	62	19	21.2	✓	70.5

All of the newly-proposed constraints have a relatively high percentage coverage and, crucially, are deemed valid by the Tolerance Principle, since the number of exceptions is in each case lower than the number of exceptions that could be tolerated. Therefore, each of the new constraints meets Enger’s (2009) first criterion.

Criterion 2: Typological parallels

The table below shows the newly-proposed constraints which have a parallel in Dutch, Norwegian and/or Spanish.

Constraint	Parallel
-EAU→*M,*F	• Nouns in -eau are neuter in Dutch e.g. <i>het cadeau, het niveau, het plateau</i>
-ET→*M,*F	• Spanish nouns in -ete are masculine e.g. <i>el filete, el sorbete</i>
-IP [i:p]→*M,*F	• Nouns with a long vowel followed by /d/ are neuter and nouns ending in a long high vowel + /l/ are feminine in Norwegian (Trosterud 2001: 47f.) • Nouns ending in -u:m, -e:m, -ym, -a:r, ie:r, -i:v and -a:t are neuter in Norwegian (Trosterud 2001: 48)

³² As discussed in 2.8.1, the Tolerance Principle technically allows all nouns to be exceptions when the total number of candidate nouns is very low. However, in this case there is no doubt that this constraint meets the first criterion, since it has no exceptions.

-AIR [ɛ:ɐ̯]→*M,*F	<ul style="list-style-type: none"> Nouns ending in <i>-ie:r</i> and <i>-a:r</i> and <i>-e:m</i> are neuter in Norwegian (Trosterud 2001: 48)
-AL ENG. →*M,*F	<ul style="list-style-type: none"> Polysyllabic nouns ending in <i>-al</i> are neuter in Norwegian (Strandskogen & Strandskogen 1986: 58)

Criterion 3: Parallels outside of gender

The features of just one of the new constraints was found to be relevant to areas of the German language other than gender.

Constraint	Parallel outside of gender assignment
-EL(L) [‘ɛl]→*M,*F	<ul style="list-style-type: none"> Adjectival suffixation: <i>-ell</i> is a “foreign” suffix commonly found on adjectives, e.g. <i>informell</i> (Duden Grammatik 2009: 752). Many of these adjectives have subsequently been converted into nouns, e.g. <i>naturell</i> > <i>das Naturell</i>. Word stress: Addition of the nominal suffix <i>-ell</i> affects word stress in that it is a stress-attracting suffix (Duden Grammatik 2009: 49).

Criterion 4: Semantic clarity

The two newly-proposed semantic constraints meet the fourth criterion in that there is extralinguistic evidence for the features they involve.

Constraint	Source of extralinguistic evidence
SYMBOLS→*M,*F	<ul style="list-style-type: none"> Wikipedia contributors. 18 June, 2016. Symbol. https://de.wikipedia.org/wiki/Symbol (10 July, 2016). Wikipedia contributors. 5 May, 2016. Sonderzeichen. https://de.wikipedia.org/wiki/Hilfe:Sonderzeichen (10 July, 2016)
RELIGIONS→*F,*N	<ul style="list-style-type: none"> Wikipedia contributors. 24 March, 2016. Liste von Religionen und Weltanschauungen. https://de.wikipedia.org/wiki/Liste_von_Religionen_und_Weltanschauungen (10 July, 2016)

Criterion 5: Accounting for exceptions

Five of the newly-proposed constraints demonstrate the ability to account for exceptions to other constraints.

Constraint	Exceptions accounted for	Total
-NNE [n(ə)]→*M,*N	FABRICS→*F,*N: <i>die Rayonne, Eolienne, Parisienne</i> MUSIC COMP/TERMS→*M,*F: <i>die Chaconne</i>	4
-EL(L) [‘ɛl]→*M,*F	GEN. PERSON→*F,*N: <i>das Modell</i> GROUPS→*M,*N: <i>das Kartell, das Artel</i>	3
-ALT→*N	CHEM.COMP.→*M,*F: <i>der Asphalt</i> SUP.→*M,*F: <i>die Anstalt</i>	2
-AL ENG.→*M,*F	MUSIC TYPES→*F,*N: <i>das Metal, der/das Spiritual</i>	2
RELIGIONS→*F,*N	-A→*M,*N: <i>der Vedanta, der Macumba</i>	2

Criterion 6: Productivity

For two of the newly-proposed constraints, there are corresponding neologisms in the OWID portal. The gender of these neologisms is in line with the constraints, thus indicating that the constraints are productive.

Constraint	Neologisms
-AL ENG. → *M, *F	<i>das Tribal</i>
SYMBOLS → *M, *F	<i>der/das Emoji, das Emoticon, das Hashtag</i>

There is also one constraint which shows diachronic productivity, in that there is historical evidence of nouns changing gender in the direction that the constraint predicts.

Constraint	Net loss/gain			Predicted direction?	Examples
	M	F	N		
-ALT → *N	+3	-1	-2	✓	<i>Halt (N>M); Gestalt (N>F)</i>

Summary of results

Below is a table summarising the findings of this section.

Constraint	1	2	3	4	5	6	TOTAL
-EL(L) [‘ɛl] → *M, *F	✓	-	✓	n/a	✓	-	3
RELIGIONS → *F, *N	✓	-	-	✓	✓	-	3
-AL ENG. → *M, *F	✓	✓	-	n/a	✓	-	3
SYMBOLS → *M, *F	✓	-	-	✓	✓	-	3
-EAU → *M, *F	✓	✓	-	n/a	-	-	2
-ET → *M, *F	✓	✓	-	n/a	-	-	2
-IP [i:p] → *M, *F	✓	✓	-	n/a	-	-	2
-NNE [n(ə)] → *M, *N	✓	-	-	n/a	✓	-	2
-AIR [ɛ:ɐ] → *M, *F	✓	✓	-	n/a	-	-	2
-ALT → *N	✓	-	-	n/a	-	✓	2

In line with the conclusions drawn in 2.9, all of the constraints proposed above can be considered valid, since they meet the first criterion, the fourth (where applicable), and at least one other. Accordingly, all of the newly-suggested constraints will be added to the constraint set, so that OGAT is better able to account for the sample data, as exemplified by (64) and (65). Some of these newly proposed constraints will be tested further in Chapter 4.

(64a)	<i>Partizip</i>	GENDER FEATURES	DEFAULT HIERARCHY		
			*N	*F	*M
	☞ <i>der Partizip</i>				*
	<i>die Partizip</i>			*!	
	<i>das Partizip</i>		*!		

(64b)	GENDER FEATURES		DEFAULT HIERARCHY		
	"STRONG"	OTHERS			
	<i>Partizip</i>	-IP [i:p]→*M,*F		*N	*F
der Partizip	*!				*
die Partizip	*!			*	
☞ das Partizip			*		

(65a)	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS	SEM. EQUIV.			
	<i>Arrival</i>		SEM. EQUIV.→*M,*N (DIE ANKUNFT)	*N	*F
der Arrival		*!			*
☞ die Arrival				*	
das Arrival		*!	*		

(65b)	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS	SEM. EQUIV.			
	<i>Arrival</i>	-AL ENG.→*M,*F	SEM. EQUIV.→*M,*N (DIE ANKUNFT)	*N	*F
der Arrival	*!	*			*
die Arrival	*!			*	
☞ das Arrival		*	*		

3.4.2 Other nouns in the sample

There is a small group of remaining nouns in the sample for which OGAT in its current state cannot account, but which can be easily accounted for if certain GENDER FEATURES constraints are ranked in relation to one another within the "other constraints" block. Two such rankings have already been proposed in this chapter, namely $-ISMA \rightarrow *M, *F \gg -A \rightarrow *M, *N$ in order to account for *Sophisma* (see 3.2.2.4), and $-ETT \rightarrow *M, *F \gg CCVC \text{ MONO} \rightarrow *F, *N$ in order to account for *Flett* (see 3.3.4.4).

The proposed rankings are presented below.

Ranking within "other constraints" block	Nouns accounted for
GROUPS→*M,*N » CVCC MONO→*F -N(C)(C) MONO→*F,*N	<i>Band</i>
FRUIT→*M,*N » CVC MONO→*F	<i>Nuss</i>
-E→*M,*N » FABRIC→*F,*N	<i>Seide</i>
-A→*M,*N » MON. UNITS→*F,*N	<i>Rupiah</i>
-NNE [n(ə)]→*M,*N » RIVERS NON-GER.→*F,*N	<i>Garonne</i>
-FT/-CHT→*M » CCVCC→*F,*N [[]C- MONO→*F,*N	<i>Schrift, Tracht</i>
ACAD. DISCIPLINES→*M,*N » CVCCC MONO→*F,*N -N(C)(C) MONO→*F,*N	<i>Kunst</i>

$SUP. \rightarrow *M, *F$ $[[[C-MONO \rightarrow *F, *N]] \gg CCVC MONO \rightarrow *F, *N$	$Kraut, (Stoff, Staat)^{33}$
--	------------------------------

With these rankings, OGAT becomes able to assign the correct genders to the nouns listed above, as exemplified by (66) and (67).

(66a)

	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS		*N	*F	*M
	$-E \rightarrow *M, *N$	FABRICS $\rightarrow *F, *N$			
<i>Seide</i>					
☞ der Seide	*				*
die Seide		*		*!	
das Seide	*	*!	*		

(66b)

	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS		*N	*F	*M
	$-E \rightarrow *M, *N$	FABRICS $\rightarrow *F, *N$			
<i>Seide</i>					
der Seide	*!				*
☞ die Seide		*		*	
das Seide	*!	*	*		

(67a)

	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS		*N	*F	*M
	$-FT/CHT \rightarrow *M$	CCVCC $\rightarrow *F, *N$			
<i>Tracht</i>					
☞ der Tracht	*				*
die Tracht		*		*!	
das Tracht		*	*!		

(67b)

	GENDER FEATURES		DEFAULT HIERARCHY		
	OTHERS		*N	*F	*M
	$-FT/CHT \rightarrow *M$	CCVCC $\rightarrow *F, *N$			
<i>Tracht</i>					
der Tracht	*!				*
☞ die Tracht		*		*	
das Tracht		*	*!		

It is important to note that no other nouns in the sample are affected (negatively) by these rankings, and that all other GENDER FEATURES constraints proposed to exist within the “other constraints” block are still postulated to be crucially equally ranked with respect to one another.

Additionally, there is evidence for the above rankings outside of the sample. For instance, the ranking $-E \rightarrow *M, *N \gg FABRIC \rightarrow *F, *N$ is supported by the existence of *die Baumwolle*, *die Viskose* and *die Chenille*;

³³ In order to account for *Kraut*, OGAT requires the ranking $SUP. \rightarrow *M, *F \gg CCVC MONO \rightarrow *F, *N$. So that this does not cause OGAT to incorrectly assign neuter to *Stoff* and *Staat*, the ranking $[[[C-MONO \rightarrow *F, *N]] \gg CCVC MONO \rightarrow *F, *N$ is required.

the ranking $-A \rightarrow *M, *N \gg \text{MON. UNITS} \rightarrow *F, *N$ is supported by *die Kuna* and *die Lira*; the ranking $-NNE [n(\partial)] \rightarrow *M, *N \gg \text{RIVERS NON-GER.} \rightarrow *F, *N$ is supported by *die Yonne*; and the ranking $-FT/-CHT \rightarrow *M \gg \text{CCVCC} \rightarrow *F, *N$; $[[j]C-\text{MONO} \rightarrow *F, *N$ is supported by nouns such as *die Schlacht* and *die Schlucht*.

It is also important to observe that the proposed rankings do not conform to any of the more general patterns of semantics \gg form (1.4.1), form \gg semantics (1.4.2), or morphology \gg semantics \gg phonology (1.4.3), but are instead individual constraint rankings irrespective of type, akin to Spitz (1965) (1.4.4).

However, many of the constraints which are proposed to be ranked below other constraints are those which relate to the phonological structure of monosyllables, such as: $\text{CCVCC} \rightarrow *F, *N$; $\text{CVCC} \rightarrow *F$ and $-N(c)(c) \text{MONO} \rightarrow *F, *N$, which brings the psychological reality of these constraints for native speakers into question. Indeed, if these constraints were deleted from the constraint set, there would be no negative impact on any of the sample nouns. Moreover, without these constraints, OGAT would be able to correctly predict the genders of *Bank* and *Macht*, which it currently cannot. OGAT would also be simplified somewhat, since it would require fewer constraints, and the need for most of the above-listed ranking arguments would be eliminated (see also 3.5.4).

Many of the proposals made in this section will be tested further in Chapter 4.

3.4.3 Testing alternative default hierarchies

As outlined in 1.4.5.1, a key component of OGAT is the default markedness hierarchy, which is proposed by Rice (2006) to comprise the ranking $*N \gg *F \gg *M$ for German. However, in 1.4.5.3, it was established that different rankings of the default markedness hierarchy could also be justified for German depending on the notion of default adopted.

For instance, on the basis of category size, or “stochastic default” (Köpcke & Zubin 2009: 148ff.), it was determined that the hierarchies $*F \gg *N \gg *M$ (supported by data from Augst 1975: 37, Köpcke 1982, Yang 1990, Schulte-Beckhausen 2002, Chan 2005 and Onysko 2007) and $*N \gg *M \gg *F$ (supported by data from Meier 1964, Rosengren 1977, Arndt 1970 and Duden Sprachwissen 2017) are also plausible (see Figures 3 and 4 in 1.4.5.3). On the basis of “system default” (Köpcke & Zubin 2009: 148ff.), it was determined that the hierarchies $*F \gg *N \gg *M$ (supported by data from Hohlfeld 2006), $*M \gg *F \gg *N$ and $*F \gg *M \gg *N$ (supported by Talanga 1987: 93, Fries 1997: 18f., Chan 2005: 75-79 and Köpcke & Zubin 2009: 148ff.) are plausible too. Finally, it was determined that on the basis of “morphological default”, the hierarchies $*N \gg *M \gg *F$ and $*M \gg *N \gg *F$ (supported by data from Hoberg 2004: 82 and Köpcke & Zubin 2009: 148ff.) are both plausible.

Consequently, these alternative possible configurations of the default hierarchy have been tested on the sample data in order to determine which ranking yields the highest number of correct predictions (i.e. fewest exceptions) and is therefore the most plausible. Note that these rankings have been tested within the revised version of OGAT which incorporates all modifications proposed in this chapter (presented in full in 3.5.1). The results are presented below.

Ranking	No. of exceptions ↓
*N»*F»*M (ORIGINAL)	30
*F»*N»*M	33
*N»*M»*F	46
*M»*N»*F	56
*F»*M»*N	56
*M»*F»*N	61

As can be seen from the table, the sample data support the original default hierarchy *N»*F»*M, as proposed by Rice (2006), since OGAT yields the most accurate results (i.e. fewest exceptions) with this configuration. The tableaux in (68) and (69) exemplify this.

(68a)

	GENDER FEATURES	DEFAULT HIERARCHY		
	OTHERS	*N	*M	*F
<i>Glockenmantel</i>				
der Glockenmantel			*!	
☞ die Glockenmantel				*
das Glockenmantel		*!		

(68b)

	GENDER FEATURES	DEFAULT HIERARCHY		
	OTHERS	*N	*F	*M
<i>Glockenmantel</i>				
☞ der Glockenmantel				*
die Glockenmantel			*!	
das Glockenmantel		*!		

(69a)

	GENDER FEATURES	DEFAULT HIERARCHY		
	OTHERS	*F	*M	*N
<i>Kühlwagen</i>				
der Kühlwagen			*!	
die Kühlwagen	*!	*		
☞ das Kühlwagen				*

(69b)

	GENDER FEATURES	DEFAULT HIERARCHY		
	OTHERS	*N	*F	*M
<i>Kühlwagen</i>				
☞ der Kühlwagen				*
die Kühlwagen	*!		*	
das Kühlwagen		*!		

*N»*F »*M is therefore the default hierarchy that will be taken forward in our revised version of OGAT. This hierarchy will be tested further in Chapter 4.

3.5 Effectiveness of OGAT vs. OGAT II and exceptions

This section will explore the overall effectiveness of OGAT in accounting for the sample data. 3.5.1 will compare the effectiveness of OGAT as originally proposed by Rice (2006) with that of a revised model which incorporates all of the adaptations proposed in this chapter. 3.5.2 will look more closely at the nouns for which the theory is unable to account. 3.5.3 will compare the effectiveness of OGAT with that of other proposed approaches to German gender assignment, and 3.5.4 will consider how the model might be made more economical.

3.5.1 OGAT vs. OGAT II

Throughout this chapter, various modifications to OGAT have been proposed in order to improve the theory’s ability to account for the sample data. These have transformed OGAT from a model in which all GENDER FEATURES constraints are crucially equally ranked with respect to one another (as illustrated in (70a)) into a model in which some GENDER FEATURES constraints are ranked higher than others (as illustrated in (70b)).

(70a)

	GENDER FEATURES	DEFAULT HIERARCHY		
		*N	*F	*M
der ___				*
die ___			*	
das ___		*		

(70b)

	GENDER FEATURES					DEFAULT HIERARCHY		
	BLOCK 1 DIMINUTIVE	BLOCK 2 “REFERENTIAL”	BLOCK 3 “STRONG”	BLOCK 4 OTHERS	BLOCK 5 SEM. EQUIV.	*N	*F	*M
der ___								*
die ___							*	
das ___						*		

In 3.3.1, it was shown that, in order to account for the genders of loanwords in the sample, OGAT requires a semantic equivalence constraint which is ranked below all other GENDER FEATURES. This appears in (70b) in BLOCK 5. In 3.3.3, it was then established that “strong” constraints are best accounted for through the ranking of these constraints above other GENDER FEATURES. Accordingly, these appear in (70b) in BLOCK 3. Subsequently in 3.3.4.1 and 3.3.4.2, it was demonstrated that constraints which determine the genders of nouns whose gender is purely semantic or “referential”

must be ranked above even “strong” GENDER FEATURES, and thus appear in BLOCK 2 in (70b), the only exception to this being constraints relating to diminutive formation, which override the assignment of “referential” gender in German and thus appear in BLOCK 1 in (70b). Finally, in 3.4.2, it was determined that in order to best account for the sample data, some rankings of individual constraints within the “other constraints” block (BLOCK 4 in (70b)) are necessary. This revised version of the theory illustrated in (70b) will henceforth be referred to as OGAT II.

As shown in the table below, OGAT in its original state as proposed by Rice (2006) is largely effective in accounting for the sample data, with the ability to account for the genders of almost 91% of nouns in the sample. However, OGAT II is able to account for a higher proportion of nouns, with a percentage coverage of almost 95%.

Approach	Sample of 592		Sample of 537	
	Exceptions	% coverage	Exceptions	% coverage
OGAT (ORIG.)	54	90.8%	50	90.7%
OGAT II	30	94.9%	28	94.8%

Whilst OGAT II might be considered to be a marginally less elegant model than the original version of OGAT, it is able to capture important generalisations about gender assignment in German which the original version cannot, as demonstrated throughout this chapter. Economy and parsimony in relation to OGAT II will be discussed further in 3.5.4.

The next subsection will consider the nouns for which OGAT II cannot account.

3.5.2 Exceptions

There are 30 nouns in the sample for which OGAT II cannot account. These are what Steinmetz (1986: 209, 2006: 1432) refers to as cases of “exceptional gender assignment” (see 1.4.5.3).

<i>das Abgangszeugnis</i>	<i>das Gebärdenspiel</i>	<i>das Kulm</i>	<i>die Stundenzahl</i>
<i>das Anderkonto</i>	<i>die Geisel</i>	<i>das Lormalphabet</i>	<i>das Uhrband</i>
<i>die Antwort</i>	<i>die Grandel</i>	<i>die Magenschleimhaut</i>	<i>die Umschau</i>
<i>der Aufzugsschacht</i>	<i>der Großbuchstabe</i>	<i>die Regel</i>	<i>die Vertretungsvollmacht</i>
<i>das Aviso</i>	<i>das Hallenbad</i>	<i>das Rektapapier</i>	<i>der Voraus</i>
<i>die C-Dur-Tonleiter</i>	<i>das Juchtenleder</i>	<i>die Rockoper</i>	<i>das Weißbuch</i>
<i>die Chargennummer</i>	<i>die Kantonalbank</i>	<i>das Schienbein</i>	
<i>das Flettnerruder</i>	<i>die Kimm</i>	<i>die Schwarzarbeit</i>	

It is widely acknowledged that all natural languages have exceptions which, at least from a synchronic perspective, appear arbitrary (Chomsky & Halle 1968: 172, Yang 2005: 266, Nübling 2011, Simon & Wiese 2011: 15). As highly complex systems which have undergone ‘diverse and interwoven historical

processes', it is to be expected that languages exhibit some degree of exceptionality (Chomsky & Halle 1968: 172, Simon & Wiese 2011: 15). Crucially, however, it is emphasised that the existence of exceptions does not undermine the existence of the remaining systematic regularities (Chomsky & Halle 1968: 172, Mel'čuk 1974: 33 cited in Corbett 1991: 8, Salmons 1993: 417, Yang 2005: 266).

While gender assignment in German has been demonstrated to be largely systematic, the genders of a small number of unaccounted-for exceptions are proposed to be memorised, i.e. stored in the lexicon alongside the noun (Köpcke & Zubin 1983: 177ff., Salmons 1993: 428, Hickey 2000: 644). As stated by Salmons (1993: 417), 'the theoretical principles of parsimony and economy would both dictate that a rule system, even one with relatively many exceptions, is far preferable to marking each noun [...] in the lexicon individually for gender'.

As outlined in 2.3, Yang (2005) maintains that it is possible to estimate the number of exceptions a system is able to tolerate before it becomes more computationally efficient to store all items individually. In this case, Yang's (2005) Tolerance Principle would permit OGAT II to have a maximum of around 92 exceptions in the sample of 592. With just 30 exceptional cases, the proposed model thus satisfies the Tolerance Principle's requirements.

It is therefore not necessary for all exceptions to be accounted for in OGAT II in order for the model to be considered plausible. Indeed, if all exceptions were accounted for, it might suggest that the model in question may in fact have been "overfit", i.e. is excessively complex and would have poor predictive abilities (e.g. Johnson 2008: 255, Smith 2011: 81, Rumshisky & Stubbs 2017: 339). Note that the predictive powers of OGAT II will be tested on a new set of data in Chapter 4.

Interestingly, Zubin & Köpcke (1981: 446f.) and Köpcke & Zubin (1997: 43f.) consider the possibility that apparent exceptions to German gender assignment might to some extent be motivated by their ability to facilitate communication. They state that, since pronouns exhibit gender agreement with their antecedent, gender is often able to disambiguate anaphoric reference, as illustrated by (71), extracted from Köpcke & Zubin (1997: 43).

(71a) *Der Krug fiel in die Schale, aber sie zerbrach nicht.*
the.MASC jug fell into the.FEM bowl but it.FEM broke not
'The jug fell into the bowl but it (the bowl) didn't break.'

(71b) *Der Krug fiel in die Schale, aber er zerbrach nicht.*
the.MASC jug fell into the.FEM bowl but it.MASC broke not
'The jug fell into the bowl but it (the jug) didn't break.'

Zubin & Köpcke (1981: 447) claim that the usefulness of gender in these contexts increases 'if there is a maximal differentiation of gender among nouns referring to items that are likely to co-occur in the

same perceptual field, or in the same text'. This explanation might therefore be used to account for cases such as: *der Löffel* 'spoon', *die Gabel* 'fork' and '*das Messer* 'knife', or *der Hals* 'throat', *die Stirn* 'forehead' and '*das Auge* 'eye'.

Consequently, it is argued that, while the principles of economy and parsimony would motivate a fully regular gender assignment system, the communicative function of gender might simultaneously motivate some degree of apparent arbitrariness within the system. However, this hypothesis would require more rigorous testing, and, furthermore, it is unclear how this explanation would account for the genders of many of the above-listed exceptions to OGAT II.

A more widely-accepted explanation for exceptionality described in the literature is frequency. Specifically, it is observed that there is a tendency for highly-frequent items to violate linguistic generalisations (Köpcke & Zubin 1983: 177ff., 1997: 47; Bybee 1988: 134; Salmons 1993: 428; Heringer 1995: 205; Nübling 2011: 139). A possible explanation for this is the claim that highly-frequent words are more likely to be stored independently, alongside all of their morphological information, and thus can be accessed without being processed. Consequently, highly-frequent words are more likely to be exceptions since by being stored they can resist regularisation (Köpcke & Zubin 1983: 177ff., 1997: 47; Bybee 1988: 134; Salmons 1993: 428). Indeed, among the above-listed exceptions to OGAT II are many nouns which are highly frequent according to the Leipzig Corpora Collection and DWDS, such as: *Spiel*, *Zahl*, *Arbeit*, *Bank*, *Buch*, *Antwort*, *Regel*, *Bad*, *Papier*, *Bein* and *Nummer*.

Furthermore, it is not only claimed that exceptions are to be expected in the highly-frequent core vocabulary, but also that exceptions are likely to occur in very peripheral, low-frequency vocabulary (Köpcke 1982, Mills 1986: 49f., Zubin & Köpcke 1984, 1986). Indeed, also among the exceptions are various context-restricted terms which have a very low frequency according to the Leipzig Corpora Collection and DWDS, namely: *der Voraus* [legal term], *das Kulm* [geological term], *die Kimm* [nautical term] and *die Grandel* [hunting term].

Der Voraus, for instance, is incorrectly predicted to be neuter by OGAT II. Interestingly, however, there is in fact evidence of *Voraus* being assigned neuter by German speakers, despite the Duden listing masculine as its only possible gender.

*„Das Voraus wird aus diesem Grunde auch als gesetzliches Vorausvermächtnis bezeichnet. [...] Erbt er aufgrund einer letztwilligen Verfügung, so erhält er **das Voraus** nicht. [...] Da der Ehegatte **das Voraus** nur erhält, wenn er gesetzlicher Erbe geworden ist [...].“* (Deutsches Anwalt Office Premium 2013)

„Das Voraus der Ehegatten“ (Boiger 2013: 177)

*„Konkurriert der Ehegatte mit Erben der 1. Ordnung, ist **das Voraus** auf die Gegenstände beschränkt [...].“* (Weißenfeld 2004: 37)

Moreover, there are in fact more Google hits for *das Voraus* than *der Voraus*: 124,000 vs. 112,000 (as of 8 September, 2017). OGAT II therefore may not be quite as “incorrect” in its prediction as some might argue. Indeed, Callies et al. (2012: 87), in their study on gender assignment to English loanwords in German, state that their findings demonstrate that native speakers often allow for more variation in grammatical gender than is reported in studies based on dictionary data. This is most likely because dictionaries have a tendency to under-report variation due to their largely prescriptive function (e.g. Davies 2012: 47).

A final observation with regard to the above-listed exceptions to OGAT II is that most of the monosyllables listed are also found by Köpcke (1982) to constitute exceptions to his proposed system. Among the monosyllables he classifies as ‘echte Ausnahmen [genuine exceptions]’ are, for instance: *Spiel, Bein, Band, Schacht* and *Haut* (Köpcke 1982: 88f., 98f.). The exceptional gender of *Kimm*, Köpcke (1982: 87) hypothesises, might be explained via a tendency for nautical terms to be feminine (although no independent evidence was found for this tendency in Chapter 2), and the exceptional genders of *Bank* and *Zahl*, Köpcke (1982: 94, 99) explains via a principle based on inflectional morphology, which, as discussed in 1.3.1.6, is unlikely to inform German gender assignment.

3.5.3 OGAT II vs. other approaches

In light of the discussion in 1.4, it is also important to consider how the effectiveness of OGAT II in accounting for the sample data compares to that of other approaches to German gender assignment. Accordingly, some of the key approaches outlined in 1.4 have been tested using the sample data, and the results have been compared with those obtained for OGAT II. The approaches tested were the semantics » form approach (e.g. Lang 1976, Corbett 1991, see 1.4.1), the form » semantics approach (see 1.4.2), the semantics » morphology » phonology approach (e.g. Köpcke 1982, see 1.4.1), and the morphology » semantics » phonology approach (e.g. Heringer 1995, Wegener 1995, see 1.4.3). The results for each approach were calculated by categorising the constraints by type (the distinction between meaning and form-based constraints is made in Appendix B) and arranging them according to the hierarchy dictated by each approach.

Approach	Sample of 592		Sample of 537	
	Exceptions	% coverage	Exceptions	% coverage
OGAT II	30	94.9%	28	94.8%
SEMANTICS » FORM	96	83.7%	89	83.4%
FORM » SEMANTICS	103	82.5%	91	83.1%
SEM » MORPH » PHON	95	83.9%	88	83.6%
MORPH » SEM » PHON	94	84.1%	85	84.2%

As illustrated above, it was found that OGAT II is considerably more successful in accounting for the sample data than the other approaches tested. This finding is significant because it challenges the claims of key gender assignment theorists such as Corbett (1991). Some specific examples of cases for which OGAT II is able to account, but some or all of the other approaches cannot, include (5)-(8), shown in 3.1.

3.5.4 Economising the theory

A final matter to consider with regard to OGAT II is economy. The idea that linguistic theories should be maximally economical is well established among generativists (as noted in e.g. Wilder & Gärtner 1997: 1ff., Radford et al. 1999: 330, Hornstein et al. 2005: 5f. etc.). Seeking the simplest and most efficient possible version of a theory is motivated by two main factors. The first is the ‘desire for formal economy’, i.e. the aim to capture generalisations with the fewest possible elements, in line with the principle of Occam’s razor (Rosenbach 2002: 235f., Davenport & Hannahs 2010: 177, Alcázar & Saltarelli 2014: 62 etc.), and the second is the value of economy and efficiency in terms of cognition and acquisition (e.g. Rosenbach 2002: 235f., Calabrese 2008: 20f., Davenport & Hannahs 2010: 177).

Economy is sought in various ways in OT, perhaps the most significant being the minimal violation of constraints by the optimal candidate, i.e. the optimal candidate is the one with fewest violations compared to other candidates (Prince & Smolensky 1993/2004: 32). In terms of the constraint set, it is stated that constraints should be simple (McCarthy 2002: 42), and that redundant constraints should not be present, for instance if their effects are replicated by the interaction of other constraints (see 2.2) (McCarthy 2002: 40, 2008: 175).

When considering economy and efficiency with regard to OGAT II, the theory can be refined in two main ways. The first is to dispense with constraints which are superfluous to the analysis, i.e. whose presence or absence does not affect the outcome of OGAT II. This strategy is also employed by e.g. Nessel (2006b: 322), who eliminates superfluous constraints in his OT analysis of Ukrainian gender. It can, however, only be determined which constraints are superfluous to OGAT II on the basis of the noun sample used in this study. It is entirely possible that a constraint which is superfluous to the current analysis may be required by nouns outside of the sample. Nevertheless, on the basis of the sample, a number of constraints might be erased for the purposes of economy and parsimony. These include: -EN→*F; STICK→*N; T- MONO→*F and D- MONO→*F.

A second way to economise OGAT II is to condense the constraint set by merging constraints which involve similar features. In his study on German, Nelson (1998: 205) advocates conflating gender assignment principles based on ‘external structural similarities or minimal phonic variants’ in order to

maximise efficiency. This idea can be extended to OGAT II, since there are a number of constraints which could be combined due to the similarity of the features they involve.

It is important to note that several of the constraints in the set already combine multiple forms based on similarity. For instance, $-HEIT/-(IG)KEIT \rightarrow *M, *N$ “nouns with the suffix *-heit* or *-(ig)keit* cannot be masculine or neuter” combines both *-heit* and *-(ig)keit* in line with the analysis that they are allomorphs of a single morpheme (e.g. Wiese 1996: 99f.). Additionally, there are constraints such as $-N(C)(C) MONO \rightarrow *F, *N$ “monosyllables whose coda consists of a nasal followed by up to two consonants cannot be feminine or neuter”, which employ formal strategies to maximise their efficiency, for example the use of parentheses, or reference to a common phonological feature such as [+NASAL] rather than to individual phonemes.

Such strategies can also be applied to a number of other constraints in the set. For instance, parentheses can be used to merge the two constraints $V_{LONG} MONO \rightarrow *F$ and $V_{LONG}C MONO \rightarrow *F$ to form the single constraint $V_{LONG}(C) MONO \rightarrow *F$. Other constraints which can be merged are those whose features are already specified by another, more general constraint. For instance, the constraints $-[I]C MONO \rightarrow *F$ and $-LT MONO \rightarrow *F$ can be readily combined, since the properties of the more specific constraint $-LT MONO \rightarrow *F$ are already included in the more general constraint $-[I]C MONO \rightarrow *F$. This is also the case for other constraints such as $-IV \rightarrow *M, *F$ and $-IV STRESSED \rightarrow *M, *F$, where the latter is subsumed under the former.

Further examples of constraints which can be conflated in order to maximise the efficiency of OGAT II are presented below.

Former separate constraints	Merged constraints
$-ANZ \rightarrow *M, *N$	$-ANZ/ENZ \rightarrow *M, *N$
$-ENZ \rightarrow *M, *N$	
$-ER SUFF. \rightarrow *F, *N$	$-ER/OR SUFF. \rightarrow *F, *N$
$-OR \rightarrow *F, *N$	
LETTERS $\rightarrow *M, *F$	LETTERS/MUS. NOTES $\rightarrow *M, *F$
MUSICAL NOTES $\rightarrow *M, *F$	
ELEMENTS $\rightarrow *M, *F$	CHEM.COMP./ELEM. $\rightarrow *M, *F$
CHEMICAL COMPOUNDS/ SUBSTANCES $\rightarrow *M, *F$	
$-INE \rightarrow *M, *N$	$-E \rightarrow *M, *N$
$-ILLE \rightarrow *M, *N$	
$-IVE \rightarrow *M, *N$	
$-E \rightarrow *M, *N$	
$-E.SUFFIX \rightarrow *M, *N$	
$-[I]C MONO \rightarrow *F$	$-L(C) MONO \rightarrow *F$
$-LT MONO \rightarrow *F$	
$-L MONO \rightarrow *F$	

It should, however, be noted that, while simplifying the constraint set is a useful way of economising the model, no limit has been set on the number of constraints permitted to exist in OT (McMahon 2000: 96) and, moreover, the measure of the success of an OT analysis is not the size of the constraint set (Ashley et al. 2010: 19). Furthermore, the number of GENDER FEATURES constraints which are proposed to exist for German³⁴ is just a fraction of the number of constraints which have been proposed for OT phonology (more than 1666 as of the year 2008) (Ashley et al. 2010). Most importantly, the number of GENDER FEATURES constraints for German is extremely small compared to the tens of thousands of nouns in the lexicon of a native speaker (e.g. Zubin & Köpcke 1981: 447), meaning that the task of learning these constraints is minimal when compared with the task of memorising the individual genders of tens of thousands of nouns. It is therefore the contention of this thesis that OGAT II is a cognitively plausible model of gender assignment.

3.6 Summary

In this chapter, the effectiveness of OGAT as a theory of German gender assignment has been tested using a set of independently verified constraints (see 2.3-2.9) and a sample of 592 nouns (see 2.2). Both the independent justification of constraints and substantial sample size are key to this investigation, since previous testing of OGAT has been based on very small data sets and has specified no limits on constraints (see 1.4.5.3).

It was found that OGAT, as proposed by Rice (2006), is relatively successful in accounting for the sample data, but that in order to account for certain subsets of nouns and to capture some important generalisations about the German gender system, a number of adaptations to the theory are required. These predominantly involve the ranking of certain groups of GENDER FEATURES constraints above others. The revised model – OGAT II – has been shown to be able to account for the genders of 95% of nouns in the sample, thereby making the model more successful in accounting for the sample data than the original version of OGAT and other key approaches to German gender assignment (outlined in 1.4).

It is now important to test the predictive power of OGAT II further, in order to determine whether the model is able to assign the correct genders when confronted with new data, which would support the idea that OGAT II models the way in which native speakers assign gender. This will be done via an

³⁴ Around 140 constraints are used in the present analysis, although some of these may prove superfluous as described above. 141 constraints remain untested (see 2.2), but it is likely that not all of these will meet Enger's (2009) criteria, and that some will prove superfluous.

experiment with native German speakers involving pseudo-nouns in Chapter 4. Note that a full summary tableau of the proposed constraint rankings within OGAT II will be presented in 4.6.

Chapter 4: Testing the predictive power of OGAT II

4.1 Introduction

In Chapter 3, an adapted version of OGAT was proposed (OGAT II) which was shown to be able to account for the genders of 95% of nouns in the sample. The proposed model involves a combination of ranking and crucial non-ranking of GENDER FEATURES constraints. In this chapter, an account will be given of an experiment that was conducted in order to further test the predictive power of this model, in which native German speakers were required to assign genders to pseudo-nouns.

Pseudo-nouns have been employed in a number of studies on gender assignment, e.g. Lang (1976), Carstensen (1980), Köpcke & Zubin (1983), Mills (1986), Wegener (1995, 2000), Schwichtenberg & Schiller (2004) and Schirrmeyer (2013) for German; Tucker et al. (1968, 1977) and Karmiloff-Smith (1979) for French; Zekhnini & Hulstijn (1995) for Dutch; Mastropavlou & Tsimpli (2011) for Greek; Sadek et al. (1975) for Spanish, as well as in various studies on other areas of nominal morphology such as plural formation, e.g. Walter (1975), MacWhinney (1978), Sonnenstuhl et al. (1999), Penke & Krause (2002).

Pseudo-nouns are effective tools for testing the psychological reality of gender assignment principles, since these items are not listed in the mental lexicon and thus the possibility of gender being assigned to them as a result of memorisation is eliminated (e.g. Mills 1986: 45, Hohlfeld 2006: 131). Experiments using pseudo-nouns therefore allow the testing of proposed mechanisms of gender assignment to determine whether or not they are used by native speakers. Moreover, pseudo-nouns can be designed with specific combinations of features which enables the testing of the interaction between particular gender assignment principles (e.g. Corbett 1991: 89).

The experiment outlined below tests a number of the proposals which were formulated in Chapter 3 on the basis of the noun sample. These regard the existence of certain constraints and the interactions between them.

4.2 Experiment design

In order to test the proposals, 26 pseudo-nouns were created. Below, each pseudo-noun is listed along with its fictitious meaning (where given), the relevant constraints to which it is subject, and an explanation of the specific proposal which is being tested.

Pseudo-noun	Meaning	Relevant constraints
<i>Zunie</i>	name of a city	CITIES/TOWNS→*M,*F -IE ['i:]/[iə]→*M,*N
<i>Yall</i>	woman's name	NAME:WOMAN→*M,*N -L(C)MONO→*F CVC MONO→*F
<i>Yällchen</i>	woman's name	-CHEN→*M,*F NAME:WOMAN→*M,*N

In 3.3.4, it was proposed that all constraints which determine the genders of nouns with so-called “referential” gender (including city/town names and personal names) are ranked above all other GENDER FEATURES constraints with the exception of -CHEN→*M,*F, which is proposed to be ranked above the constraints which determine the genders of nouns with referential gender.

If this is the case, then the constraint CITIES/TOWNS→*M,*F will be ranked above -IE ['i:]/[iə]→*M,*N (a “strong” constraint which otherwise categorically determines feminine gender), and *Zunie* will be assigned neuter. Additionally, the constraint NAME:WOMAN→*M,*N will be ranked above both -L(C)MONO→*F and CVC MONO→*F (if it indeed exists, see 3.4.2), and *Yall* will be assigned feminine. Furthermore, -CHEN→*M,*F will be ranked above NAME:WOMAN→*M,*N, meaning that *Yällchen* will be assigned neuter.

Pseudo-noun	Meaning	Relevant constraints
<i>Springling</i>	a flower	-LING→*F,*N FLOWERS→*M,*N
<i>Randor</i>	a tree	-OR→*F,*N TREES→*M,*N
<i>Mellade</i>	a fish	-ADE→*M,*N FISH→*F,*N

In 3.3.3, it was proposed that all “strong” constraints, i.e. those involving features which usually categorically determine gender (except for in cases of referential gender), e.g. certain suffixes, are ranked above all other GENDER FEATURES constraints (except for those determining referential gender as described above).

If this is the case, -LING→*F,*N will be ranked above FLOWERS→*M,*N, causing *Springling* to be assigned masculine, and -OR→*F,*N will be ranked above TREES→*M,*N, meaning that *Randor* will also be assigned masculine. In addition, -ADE→*M,*N will be ranked above FISH→*F,*N, resulting in *Mellade* being assigned feminine.

Pseudo-noun	Meaning	Relevant constraints
<i>Syllenage</i>	a metal	METALS→*M,*F -AGE→*M,*N
<i>Vanade</i>	a language	LANGUAGES→*M,*F -ADE→*M,*N
<i>Gamie</i>	a letter of the alphabet	LETTERS/MUS. NOTES→*M,*F -IE ['i:]/[iə]→*M,*N
<i>Arlt</i>	type of cheese	CHEESES→*F,*N VCCC MONO→*M -L(C) MONO→*F

As stated above, it is proposed in 3.3.3 that “strong” constraints are ranked above all other GENDER FEATURES constraints (except for those determining referential gender). It is, however, unclear whether all such constraints (e.g. METALS→*M,*F and AGE→*M,*N) are crucially equally ranked with respect to one another, or whether a hierarchical relationship exists between them. In 3.3.4.2, it was proposed that the semantic constraints which usually categorically determine gender (e.g. METALS→*M,*F; LANGUAGES→*M,*F) might be ranked alongside the constraints determining referential gender (e.g. CITIES/TOWNS→*M,*F) and therefore above the otherwise categorical form-based constraints such as -ADE→*M,*N and IE ['i:]/[iə]→*M,*N. If this is the case, then *Syllenage*, *Vanade* and *Gamie* would be assigned neuter, and *Arlt* would be assigned masculine.

Pseudo-noun	Meaning	Relevant constraints
<i>Zabe</i>	a type of fabric	-E→*M,*N FABRIC→*F,*N
<i>Schnucht</i>	---	-FT/-CHT→*M [[]C- MONO→*F,*N CCVCC→*F,*N
<i>Bazka</i>	a currency	MON. UNITS→*F,*N -A→*M,*N
<i>Aranne</i>	a French river	-NNE [N(ə)]→*M,*N RIVERS NON-GER.→*F,*N RIVERS EUROPE→*N

In 3.4.2, it was proposed that -E→*M,*N is ranked above FABRIC→*F,*N on the basis of *die Seide*, and that -FT/-CHT→*M is ranked above [[]C- MONO→*F,*N (and CCVCC→*F,*N if it exists, see 3.4.2) on the basis of *die Schrift*. It was also proposed that -A→*M,*N is ranked above MON. UNITS→*F,*N on the basis of *Rupiah*. Additionally, in 3.4.1, the constraint -NNE [N(ə)]→*M,*N was proposed, and in 3.4.2 this was proposed to be ranked above RIVERS NON-GER.→*F,*N on the basis of *die Garonne*. If each of these rankings are indeed accurate, then it is predicted that *Zabe*, *Aranne*, *Schnucht* and *Bazka* will all be assigned feminine.

Pseudo-noun	Meaning	Relevant constraints
<i>Ahb</i>	an alcoholic drink	ALC. DRINKS→*F,*N VC MONO→*M,*F V _{LONG} (C) MONO→*F
<i>Fralang</i>	a type of dwelling	-ANG→*F,*N DWELL→*M

Other than those specifically discussed in Chapter 3, all other GENDER FEATURES constraints are hypothesised to be crucially equally ranked with respect to one another (as per Rice's (2006) original proposal). If this is indeed the case, and if the default hierarchy *N»*F»*M is correct (see also 3.4.3), then both *Ahb* and *Fralang* will be assigned masculine.

Pseudo-noun	Meaning	Relevant constraints
<i>Placht</i>	---	-FT/-CHT→*M CCVCC→*F,*N
<i>Grett</i>	---	-ETT→*M,*F CCVC MONO→*F,*N
<i>Jaff</i>	a type of fruit	FRUIT→*M,*N CVC MONO→*F

In 3.4.2, it was proposed that the constraint CCVCC→*F,*N either does not exist or that it is ranked below -FT/-CHT→*M on the basis of the nouns in the sample *die Tracht* and *die Schrift*. If either of these statements is true, *Placht* will be assigned feminine. Similarly, it was proposed in 3.3.4.4 and 3.4.2 that the constraint CCVC MONO→*F,*N either does not exist or that it is ranked below ETT→*M,*F on the basis of the noun in the sample *das Flett*. If either is true, *Grett* will be assigned neuter. It was also hypothesised in 3.4.2 that the constraint CVC MONO→*F is either not a psychologically real constraint or, if it is, that it is ranked below FRUIT→*M,*N on the basis of the noun in the sample *die Nuss*. If either of these proposals is true, *Jaff* will be assigned feminine.

Pseudo-noun	Meaning	Relevant constraints
<i>Felpt</i>	a game	GAMES→*M,*F CVCCC MONO→*F,*N
<i>Henn</i>	a scientific unit	SCI. UNITS→*M,*F -N(C)(C) MONO→*F,*N CVC MONO→*F
<i>Lasp</i>	an insect	INSECTS→*M,*N CVCC MONO→*F

If both GAMES→*M,*F and CVCCC MONO→*F,*N are constraints in existence in the gender assignment system of a native German speaker, OGAT II would predict *Felpt* to be masculine. However, on the basis of nouns in the sample such as *die Kunst*, it was hypothesised in 3.4.2 that CVCCC MONO→*F,*N may not be a psychologically real constraint for German native speakers. If it is not, then *Felpt* should be assigned neuter.

In 3.4.2, it was also questioned whether the constraint $-N(C)(C) \text{ MONO} \rightarrow *F, *N$ is a psychologically real constraint, based on nouns in the sample such as *die Kunst, die Band, die Kimm, die Bank* etc. If it is not, then *Henn* is predicted to be neuter. Similarly, the existence of the constraint $\text{CVCC MONO} \rightarrow *F$ was questioned based on nouns in the sample such as *die Band, die Bank, die Macht* etc. If it does not exist, then *Lasp* is predicted to be feminine.

Pseudo-noun	Meaning	Relevant constraints
<i>Klepisma</i>	---	-ISMA $\rightarrow *M, *F$ -A $\rightarrow *M, *N$

In 3.2.2.4, the constraint $-\text{ISMA} \rightarrow *M, *F$ was proposed, and it was hypothesised that it exists in a position that is ranked above the constraint $-\text{A} \rightarrow *M, *N$ on the basis of the noun in the sample *das Sophisma*. If this is indeed the case, *Klepisma* will be assigned neuter.

Pseudo-noun	Meaning	Relevant constraints
<i>Grileau</i>	---	-EAU $\rightarrow *M, *F$
<i>Trelair</i>	---	-AIR [ɛ:ɛ̃] $\rightarrow *M, *F$
<i>Dranalt</i>	---	-ALT $\rightarrow *N$

Finally, in 3.4.1, a number of other new constraints were proposed on the basis of several nouns in the sample. These include: $-\text{EAU} \rightarrow *M, *F$ (proposed on the basis of *das Rondeau*); $-\text{AIR} [\epsilon:\epsiloñ] \rightarrow *M, *F$ (proposed on the basis of *das Eclair*); and $-\text{ALT} \rightarrow *N$ (proposed on the basis of *der Asphalt* and *die Anstalt*). If $-\text{EAU} \rightarrow *M, *F$ and $-\text{AIR} [\epsilon:\epsiloñ] \rightarrow *M, *F$ are indeed valid constraints, then it is expected that *Grileau* and *Trelair* will be assigned neuter. If $-\text{ALT} \rightarrow *N$ is a valid constraint, and if the default hierarchy $*N \gg *F \gg *M$ is correct, *Dranalt* will be assigned masculine.

4.3 Method

Participants

100 adult native German speakers took part in the experiment. This is considerably more than in the majority of previous studies using pseudo-nouns to investigate gender assignment (e.g. Carstensen 1980, Köpcke & Zubin 1983, Mills 1986, Wegener 1995, Schwichtenberg & Schiller 2004, Hohlfeld 2006, Mastropavlou & Tsimpli 2011, Schirrmeyer 2013). Their participation in the study was voluntary. All participants were from Germany and were asked to specify their region of origin.

Materials

An online questionnaire was created which required the participants to assign a gender to the 26 above-listed pseudo-nouns. This number of test items was chosen as it allows the testing of a range of proposals whilst remaining small enough to ensure questionnaire completion. Exactly the same number of test items was used in e.g. Callies et al. (2010, 2012).

All pseudo-nouns were designed to be phonotactically permissible in German and to have one or more features relevant to gender assignment so that the proposals listed above could be tested. In order to ensure that the pseudo-nouns (or any highly similar nouns – see below) did not already exist, the words were checked with native speakers, online search engines and dictionaries.

In cases where a semantic constraint was relevant to the investigation, the pseudo-nouns were given a fictitious meaning. For these pseudo-nouns, an example sentence was provided in order to, firstly, clarify the meaning of the item (cf. Fischer 2005: 284f.), and secondly, to provide a context for the insertion of a definite article (as done in e.g. Lang 1976; Callies et al. 2010, 2012; Franco et al. 2018). All definite articles to be inserted were in the nominative singular.

The order of the questions was randomised for each participant so that any possible sequence effects could be avoided. Multiple orderings were also used in e.g. MacWhinney (1978), Köpcke & Zubin (1983), Mills (1986), Wegener (1995), Callies et al. (2010, 2012) and Franco et al. (2018) for this reason.

An important consideration was to minimise the possibility that the participants simply assign gender to the pseudo-nouns by means of analogy with existing German nouns (a possibility also considered by Köpcke & Zubin (1983), MacWhinney (1978), Mills (1986), Wegener (1995) and Hohlfeld (2006)). Following a similar method to Köpcke & Zubin (1983), a pre-experiment test was carried out in which three native German speakers were presented with the 26 pseudo-nouns and asked to spontaneously produce the first similar-sounding word that came to mind. Any pseudo-noun which triggered three identical responses would be changed. The only pseudo-noun for which this was the case was *Schrucht*, where all three native speakers produced *Schlucht*. Consequently, this was changed to *Schnucht*, which did not trigger three identical responses.

Procedure

The participants were firstly shown a trial question to illustrate the format of the experiment, as was also done in e.g. Lang (1976), Carstensen (1980), Mills (1986), Schwichtenberg & Schiller (2004) and Hohlfeld (2006). They were then shown each of the 26 pseudo-nouns in turn (in a written format) and were asked to select a definite article (*der*, *die* or *das*) for each one. The participants were able to

select more than one article if they deemed it necessary to do so. They also had the option of adding a comment to their answer, should they have wished to justify their response. While any justifications should not, of course, be taken as a necessarily valid explanation for the results obtained, it is nonetheless interesting to observe which features participants consider to be relevant to a noun's gender.

There was no time limit (as with the majority of similar studies), but the participants were told in advance that completion of the questionnaire would take approximately five to ten minutes. Additionally, participants were explicitly instructed to answer with their instinctive reaction and to not think too long about their answers (as also done in e.g. Callies et al. 2010, 2012). They were also told that there were no correct or incorrect answers in order to dissuade them from seeking guidance from reference works. The questionnaire platform automatically recorded the time taken by each participant to complete the questionnaire.

The questionnaire can be seen in full in Appendix G.

Hypotheses

Null hypothesis (H_0): The semantic, morphological and phonological features of the pseudo-nouns will not influence the gender assigned to them by native German speakers. That is to say, the difference between the probability of native speakers selecting a gender for each pseudo-noun at random (33.3% for each gender)¹ and the actual distribution of results will not be statistically significant ($p = >0.05$).

Alternative hypothesis (H_a): The semantic, morphological and phonological features of the pseudo-nouns will influence the gender assigned to them by native German speakers (specific proposals outlined in 4.2). That is to say, the difference between an entirely random distribution across the three genders for each pseudo-noun and the actual distribution of results will be statistically significant ($p = \leq 0.05$).

¹ As also done by e.g. Lang (1976), Mills (1986), Wegener (1995) and Hohlfeld (2006), this statistic has been calculated on the basis that there are three possible genders for each pseudo-noun. However, it might also be argued that a random distribution of results should instead reflect the overall distribution of genders in the lexicon, as done by e.g. Schwichtenberg & Schiller (2004), who use a calculation based on the token frequency of 4113 simplex nouns in the CELEX database (Baayen et al. 1995). Nevertheless, as demonstrated in 1.4.5.3, estimates of the distribution of genders in the lexicon vary greatly depending on the source material and thus a calculation of random distribution based on one of these estimates would be unreliable. Moreover, the calculation used by Schwichtenberg & Schiller (2004) (39:35:26) does not deviate significantly from the 33:33:33 distribution used in the present study ($p = 0.3$).

4.4 Results

All 100 participants answered all 26 questions, and the total number of responses for each category was calculated for each pseudo-noun. According to the questionnaire platform, the median time taken to complete the survey was 00:06:41, which yields an average of 13.8 seconds per question, assuming that approximately the same amount of time was dedicated to each survey slide (including an introductory slide, an instruction slide and a trial question). Such a short median response time suggests that at least most of the participants did indeed give their instinctive reaction for each question and did not consult any reference works.

In order to perform a statistical analysis on the results obtained, it was firstly necessary to divide any answers in which participants had selected more than one gender (highlighted in grey below) by the number of genders they had selected and add the resulting figure to the totals of the relevant gender categories. For example, if a participant had selected both masculine and neuter for a pseudo-noun, then this result would have been divided into two, contributing 0.5 to the total for masculine and 0.5 to the total for neuter. This then provides us with three totals (one for each gender) on which to perform the statistical analyses. The proportion of multiple-gender answers was relatively low (8.2% of total responses) and will be discussed below in further detail for each pseudo-noun where relevant.

Pseudo-noun	M	F	N	M/N	F/N	M/F	M/F/N	M	F	N
<i>Arlt</i>	94	3	1	2	0	0	0	95.0	3.0	2.0
<i>Mellade</i>	2	92	0	0	0	6	0	5.0	95.0	0.0
<i>Aranne</i>	5	92	1	0	0	2	0	6.0	93.0	1.0
<i>Randor</i>	88	6	0	0	0	6	0	91.0	9.0	0.0
<i>Springling</i>	86	7	2	3	1	1	0	88.0	8.0	4.0
<i>Yall</i>	12	81	1	2	0	4	0	15.0	83.0	2.0
<i>Schnucht</i>	10	79	5	0	1	5	0	12.5	82.0	5.5
<i>Felpt</i>	13	1	78	7	1	0	0	16.5	1.5	82.0
<i>Yällchen</i>	2	15	78	0	5	0	0	2.0	17.5	80.5
<i>Placht</i>	13	73	4	1	1	5	3	17.0	77.0	6.0
<i>Dranalt</i>	69	8	11	4	0	5	3	74.5	11.5	14
<i>Klepisma</i>	5	17	68	1	5	1	3	7.0	21.0	72.0
<i>Henn</i>	19	9	67	2	2	1	0	20.5	10.5	69.0
<i>Gamie</i>	18	12	63	4	2	1	0	20.5	13.5	66.0
<i>Zunie</i>	2	29	65	0	2	2	0	3.0	31.0	66.0
<i>Grett</i>	17	7	57	13	1	2	3	25.5	9.5	65.0
<i>Fralang</i>	60	4	26	8	1	0	1	64.3	4.8	30.8
<i>Ahb</i>	56	0	32	7	0	3	2	61.7	2.2	36.2
<i>Trelair</i>	20	12	55	7	2	1	3	25.0	14.5	60.5
<i>Grileau</i>	28	6	54	8	1	0	3	33.0	7.5	59.5
<i>Syllenage</i>	2	57	34	2	5	0	0	3.0	59.5	37.5
<i>Vanade</i>	0	38	55	0	6	0	1	0.3	41.3	58.3
<i>Bazka</i>	51	36	3	0	0	9	1	55.8	40.8	3.3
<i>Jaff</i>	28	51	11	4	2	1	3	31.5	53.5	15.0
<i>Zabe</i>	27	49	15	1	4	3	1	29.3	52.8	17.8
<i>Lasp</i>	31	49	12	0	2	5	1	33.8	52.8	13.3

Figure 1: Division of multiple-gender answers into the three gender categories

The first statistical analysis performed on the results was a chi-square (χ^2) test which was used to assess whether the actual distribution of results obtained for each gender is significantly different from a random distribution (33:33:33) that would be expected if the null-hypothesis were true. The results of this test are given for each pseudo-noun in the table below, ordered from the highest χ^2 statistic (i.e. lowest p-value) to the lowest χ^2 statistic (i.e. highest p-value).

Pseudo-noun	M	F	N	χ^2 statistic ↓	p-value
<i>Mellade</i>	5.0	95.0	0.0	171.67	<0.001***
<i>Arlt</i>	95.0	3.0	2.0	171.31	<0.001***
<i>Aranne</i>	6.0	93.0	1.0	160.74	<0.001***
<i>Randor</i>	91.0	9.0	0.0	151.01	<0.001***
<i>Springling</i>	88.0	8.0	4.0	134.85	<0.001***
<i>Yall</i>	15.0	83.0	2.0	113.65	<0.001***
<i>Felpt</i>	16.5	1.5	82.0	110.07	<0.001***
<i>Schnucht</i>	12.5	82.0	5.5	107.42	<0.001***
<i>Yällchen</i>	2.0	17.5	80.5	103.82	<0.001***
<i>Placht</i>	17.0	77.0	6.0	87.71	<0.001***
<i>Dranalt</i>	74.5	11.5	14.0	76.43	<0.001***
<i>Klepisma</i>	7.0	21.0	72.0	70.29	<0.001***
<i>Zunie</i>	3.0	31.0	66.0	59.84	<0.001***
<i>Henn</i>	20.5	10.5	69.0	58.80	<0.001***
<i>Ahb</i>	61.7	2.2	36.2	53.52	<0.001***
<i>Fralang</i>	64.3	4.8	30.8	53.44	<0.001***
<i>Vanade</i>	0.3	41.3	58.3	53.39	<0.001***
<i>Grett</i>	25.5	9.5	65.0	49.01	<0.001***
<i>Gamie</i>	20.5	13.5	66.0	48.80	<0.001***
<i>Syllenage</i>	3.0	59.5	37.5	48.71	<0.001***
<i>Bazka</i>	55.8	40.8	3.3	43.92	<0.001***
<i>Grileau</i>	33.0	7.5	59.5	40.61	<0.001***
<i>Trelair</i>	25.0	14.5	60.5	34.90	<0.001***
<i>Lasp</i>	33.8	52.8	13.3	23.44	<0.001***
<i>Jaff</i>	31.5	53.5	15.0	22.41	<0.001***
<i>Zabe</i>	29.3	52.8	17.8	19.11	<0.001***

Figure 2: The significance of the distribution of results obtained

The difference between the distribution of results obtained and a random distribution across the genders was shown to be highly significant ($p < 0.001$) for all pseudo-nouns.² While a lower χ^2 statistic was calculated for the pseudo-nouns towards the bottom of the table, showing a slightly more even distribution of results across the three genders than those listed towards the top of the table, the p-

² Assuming a significance level of 0.05, as is standard in linguistics (e.g. Baayen 2008: 68). This is also the significance level used by e.g. Schwichtenberg & Schiller (2004) and Hohlfeld (2006).

value is below 0.001 for all test items. This means that there is less than a 0.1% chance that the actual distribution of results was achieved at random, which constitutes strong evidence against the null-hypothesis. That is to say, the data demonstrate that it is extremely unlikely that the features of the pseudo-nouns did not influence the gender assigned to them by the participants.

In order to evaluate the validity of the specific proposals outlined in 4.2, however, it is necessary to perform a further statistical test. While it has been shown that the distribution of results is significant, i.e. the participants did not assign the genders in equal measure for any of the pseudo-nouns, it is now necessary to determine whether the results obtained for the most commonly selected gender are significantly higher than those obtained for the gender which was selected the second most often. For instance, a pseudo-noun yielding the result 0:49:51 across the three genders has a distribution which is significantly different from 33:33:33 ($p < 0.001$), yet the difference between the most commonly selected gender and the second most commonly selected gender is not significant ($p = 0.84$), thus it cannot be said that the gender with the highest result was selected significantly most often.

Accordingly, a further chi-square test was used to determine whether the distribution of the results obtained for the two most commonly selected genders was significantly different from an even distribution between the two. The results of this test are shown in Figure 3 below.

Pseudo-noun	Highest result	Second highest result	χ^2 statistic ↓	p-value
<i>Arlt</i>	95.0 (M)	3.0 (F)	86.37	<0.001***
<i>Mellade</i>	95.0 (F)	5.0 (M)	81.00	<0.001***
<i>Aranne</i>	93.0 (F)	6.0 (M)	76.45	<0.001***
<i>Randor</i>	91.0 (M)	9.0 (F)	67.24	<0.001***
<i>Springling</i>	88.0 (M)	8.0 (F)	66.67	<0.001***
<i>Yall</i>	83.0 (F)	15.0 (M)	47.18	<0.001***
<i>Schnucht</i>	82.0 (F)	12.5 (M)	51.11	<0.001***
<i>Felpt</i>	82.0 (N)	16.5 (M)	43.56	<0.001***
<i>Dranalt</i>	74.5 (M)	14.0 (N)	41.36	<0.001***
<i>Yällchen</i>	80.5 (N)	17.5 (F)	40.50	<0.001***
<i>Placht</i>	77.0 (F)	17.0 (M)	38.30	<0.001***
<i>Klepisma</i>	72.0 (N)	21.0 (F)	27.97	<0.001***
<i>Henn</i>	69.0 (N)	20.5 (M)	26.28	<0.001***
<i>Gamie</i>	66.0 (N)	20.5 (M)	23.93	<0.001***
<i>Zunie</i>	66.0 (N)	31.0 (F)	12.63	<0.001***
<i>Fralang</i>	64.3 (M)	30.8 (N)	11.80	<0.001***
<i>Trelair</i>	60.5 (N)	25.0 (M)	14.74	<0.001***
<i>Grett</i>	65.0 (N)	25.5 (M)	17.24	<0.001***
<i>Grileau</i>	59.5 (N)	33.0 (M)	7.59	<0.01**
<i>Zabe</i>	52.8 (F)	29.3 (M)	6.73	<0.01**
<i>Ahb</i>	61.7 (M)	36.2 (N)	6.64	<0.01**

<i>Jaff</i>	53.5 (F)	31.5 (M)	5.69	<0.02**
<i>Syllenage</i>	59.5 (F)	37.5 (N)	4.99	<0.05*
<i>Lasp</i>	52.8 (F)	33.8 (M)	4.17	<0.05*
<i>Vanade</i>	58.3 (N)	41.3 (F)	2.90	<0.1
<i>Bazka</i>	55.8 (M)	40.8 (F)	2.44	<0.2

Figure 3: The significance of the difference between the highest and second highest result

For all but two pseudo-nouns, the difference between the actual distribution of results for the two most commonly selected genders and an even distribution is statistically significant given a significance level of 0.05. For the first 18 pseudo-nouns listed in Figure 3, this difference was highly significant ($p < 0.001$), for four pseudo-nouns (*Grileau*, *Zabe*, *Ahb* and *Jaff*) the difference was moderately significant ($p < 0.01$, < 0.02) and for two pseudo-nouns (*Syllenage* and *Lasp*) the difference was weakly significant ($p < 0.05$). The difference was found to be insignificant for just two of the test items: *Vanade* ($p < 0.1$) and *Bazka*, ($p < 0.2$). It can therefore be concluded that the gender with the highest result was selected significantly most often for 24 of the pseudo-nouns. The implications of this conclusion on the specific proposals outlined in 4.2 will be discussed in the following sections.

4.5 Discussion

In this section, the above results will be discussed in terms of each individual pseudo-noun.

***Zunie* (city): 2(M), 29(F), 65(N), 2(F/N), 2(M/F)**

The responses for this pseudo-noun are almost entirely distributed between feminine and neuter, with only 4 participants considering masculine as a possible gender. The selection of feminine and neuter as the two most plausible options demonstrates the presence of the two constraints $-IE [i:] / [iə] \rightarrow *M, *N$ and $CITIES/TOWNS \rightarrow *M, *F$. As shown in 4.4, the proportion of participants selecting neuter is greater to a high level of significance than those selecting feminine ($p < 0.001$), confirming the hypothesis made in 3.3.4.1 that $CITIES/TOWNS \rightarrow *M, *F$ is ranked above $-IE [i:] / [iə] \rightarrow *M, *N$, and contributing evidence to support the claim that “strong” constraints which otherwise categorically determine gender are ranked below those which determine the gender of nouns with referential gender.

Although the number of participants selecting neuter is statistically significantly higher than the number selecting feminine, it is perhaps not as high as might be expected. One possible reason for this could be the fact that cities and towns are used with gender-marked targets in only few contexts in German (e.g. Duden Grammatik 2009: 299ff.). Despite the example sentence providing one of these

few contexts (*D___ heutige Zunie ist mit 1,6 Millionen Einwohnern die zweitgrößte Stadt des Landes*), two participants stated in the comments section of their answer that they would either prefer not to use an article with the pseudo-noun in that context (“*eigtl. kein Artikel*” (response: M/F)) or that they would consider it possible to use the pseudo-noun without an article in that context (“*oder kein Artikel*” (response: F)). This suggests that at least some native speakers might not usually use or hear gender-marked targets with city/town names, meaning that the constraint CITIES/TOWNS→*M,*F may not be reinforced very often, and thus some speakers might not be aware of the fact that cities and towns are standardly used with neuter targets. Such speakers may have assigned *Zunie* masculine based on the application of a constraint CITIES/TOWNS→*F,*N, or assigned it feminine simply based on its form (i.e. applying only the constraint -IE [‘i:’]/[‘iə’]→*M,*N).

Another possible reason for the relatively high proportion of feminine assignment to *Zunie* is highlighted by Carstensen (1980: 20), who, based on the results of his study of gender assignment to pseudo-nouns, hypothesises that some speakers may assign gender to a pseudo-noun based on its hypernym as given in the definition. In the case of *Zunie*, this would mean that some speakers might have assigned the pseudo-noun feminine based on its feminine hypernym *Stadt* ‘city’ as given in the definition ‘*Stadt in Burkina-Faso*’.

Yall (woman’s name): 12(M), 81(F), 1(N), 2(M/N), 4(M/F)

The vast majority of participants assigned feminine to *Yall*, providing evidence to support the proposal that constraints which determine the gender of nouns with referential gender, e.g. NAME:WOMAN→*M,*N, are ranked above all other constraints (apart from -CHEN→*M,*F), e.g. -L(C) MONO→*F and CVC MONO→*F. The difference between the proportion of participants assigning feminine to *Yall* and the proportion assigning the second most popular gender, masculine, is highly significant ($p < 0.001$).

One possible reason for 18 participants considering masculine to be a possible gender for *Yall* is that, while the definition stated that the name is a ‘*weiblicher Vorname* [woman’s first name]’, the example sentence given did not explicitly indicate that the referent of *Yall* was a woman (*Da kommt d___ kluge Yall von gegenüber*). Therefore, it is possible that some participants focussed only on the example sentence and assumed that *Yall* was referring to a man, or considered it possible that *Yall* could be referring to a man in this context (thus applying the constraint NAME:MAN→*F,*N).

The three participants who considered neuter to be a possible gender for *Yall* may have done so due to regional differences (hence all participants being asked to specify their region of origin). The three participants in question originate from South-West and West-Central Germany, which are regions

where neuter targets can be used with nouns referring to women (e.g. Elspaß & Möller 2003c, Nübling et al. 2013, see also 3.3.4.2). Therefore, for speakers in these regions, the constraint would instead be NAME:WOMAN→*M,*F, causing *Yall* to be assigned neuter.

***Yällchen* (woman's name): 2(M), 15(F), 78(N), 5(F/N)**

For *Yällchen*, the gender selected most often by the participants was neuter, thus supporting the proposal that the constraint -CHEN→*M,*F is ranked above constraints determining the gender of nouns with referential gender such as NAME:WOMAN→*M,*N. The second most commonly-selected gender (albeit comparatively insignificantly so) was feminine, which is most plausibly explained by speakers analysing *Yällchen* as simplex. While many of the participants did analyse *-chen* as a diminutive morpheme, demonstrated by several comments, e.g. “*Diminutiv = das*”, “*Verniedlichung = neutrum*”, “*Verkleinerungsform*” (all selecting neuter), others expressed uncertainty of the status of *-chen* in this pseudo-noun, leaving comments such as: “*das -chen suggeriert für mich Diminutiv, wenn es sich wirklich um einen Namen handeln soll, würde ich immer die nehmen* [the *-chen* suggests a diminutive to me, if it is really supposed to be a name, I would always choose *die*]” (response: F/N). This further confirms the conclusion of the discussion in 3.3.4.2 that *-chen* is only a neuter-assigning feature when it is analysed as a diminutive suffix.

***Springling* (flower): 86(M), 7(F), 2(N), 3(M/N), 1(F/N), 1(M/F)**

The vast majority of participants assigned masculine to *Springling*, which is in line with the hypothesis that -LING→*F,*N, as a constraint which categorically determines the gender of nouns (except for in cases of referential gender), is ranked above FLOWERS→*M,*N. However, it is important to note that the same result would also have been achieved if both of these constraints were crucially equally ranked, due to the default markedness hierarchy. The second highest result, which was very much statistically significantly lower than the highest result ($p < 0.001$), was feminine, which demonstrates the form/meaning conflict in the assignment of gender to this pseudo-noun.

***Randor* (tree): 88(M), 6(F), 0(N), 6(M/F)**

A comparable result was achieved for *Randor*, i.e. a significantly high proportion of participants selecting masculine, which is similarly in line with the hypothesis that -OR→*F,*N, as a “strong” constraint, is ranked above TREES→*M,*N. However, an alternative explanation for this result could also be the crucial equal ranking of these constraints. The only non-masculine results obtained for *Randor* were those assigning feminine to the pseudo-noun, and six participants selecting both feminine and masculine, which again demonstrates the conflict between meaning and form.

***Mellade* (fish): 2(M), 92(F), 0(N), 6(M/F)**

A significantly high proportion of participants assigned feminine to *Mellade*, which is consistent with the proposal that $-ADE \rightarrow *M, *N$, as a “strong” constraint, is ranked above $FISH \rightarrow *F, *N$. Unlike for *Springling* and *Randor*, the same result could not have been achieved via the crucial equal ranking of these constraints. One alternative explanation for the results could be the non-existence of the constraint $FISH \rightarrow *F, *N$, however there is no independent reason to suggest this. The (very few) non-feminine results for *Mellade* were for masculine, and six participants selected both feminine and masculine, which again illustrates the form/meaning conflict for this pseudo-noun.

***Gamie* (letter): 18(M), 12(F), 63(N), 4(M/N), 2(F/N), 1(M/F)**

The results show a statistically significant proportion of results for neuter, supporting the hypothesis that $LETTERS/MUS. NOTES \rightarrow *M, *F$ is ranked alongside the constraints which determine the gender of nouns with referential gender and thus above “strong” constraints such as $-E [i:] / [i\partial] \rightarrow *M, *N$.

Surprisingly, the second highest result was obtained for masculine, and not feminine as might have been expected due to the pseudo-noun’s form. This might be explained, firstly, by Carstensen’s (1980: 20) hypernym theory, given that the hypernym specified in the definition was *der Buchstabe* ‘letter’, or secondly, by the fact that, as with city names, gender-marked targets are used with letters only infrequently, thus causing some speakers to be unsure of the gender that is standardly used with such nouns. This latter explanation is supported by one comment left by a participant, stating that they would prefer not to use an article with *Gamie*: “*eigtl. kein Artikel*” (response: M).

***Arlt* (cheese): 94(M), 3(F), 1(N), 2(M/N)**

Almost all participants assigned masculine to *Arlt*, which is consistent with the proposal that $CHEESES \rightarrow *F, *N$ is ranked alongside the constraints which determine the gender of nouns with referential gender and thus above constraints which would otherwise usually categorically determine gender such as $VCCC MONO \rightarrow *M$. An alternative explanation for the results obtained could, however, be that these two constraints are crucially equally ranked. It is also possible that the constraint $VCCC MONO \rightarrow *M$ is not a psychologically real constraint in the gender assignment system of a native German speaker, or that neither $VCCC MONO \rightarrow *M$ nor $CHEESES \rightarrow *F, *N$ exist as constraints, and *Arlt* is assigned masculine by default, however there is no independent reason to suggest this. Six participants considered non-masculine genders as possible responses, which demonstrates the conflict between $VCCC MONO \rightarrow *M$ and $CHEESES \rightarrow *F, *N$.

Vanade (language): 0(M), 38(F), 55(N), 6(F/N), 1(M/F/N)

The results for *Vanade* are almost entirely distributed between feminine and neuter, which is expected due to the presence of the two constraints $\text{LANGUAGES} \rightarrow *M, *F$ and $-\text{ADE} \rightarrow *M, *N$. Despite the majority of participants selecting neuter for this pseudo-noun, the difference between the number selecting neuter and the number selecting feminine is not statistically significant ($p < 0.1$) given a significance level of 0.05.

Therefore, while we might accept the proposal that $\text{LANGUAGES} \rightarrow *M, *F$ is ranked alongside the constraints determining the gender of nouns with referential gender, and thus above $-\text{ADE} \rightarrow *M, *N$, the statistical evidence to support this is inconclusive. The possibility that $\text{LANGUAGES} \rightarrow *M, *F$ and $-\text{ADE} \rightarrow *M, *N$ are crucially equally ranked therefore remains.

The ranking of $\text{LANGUAGES} \rightarrow *M, *F$ above $-\text{ADE} \rightarrow *M, *N$ is, however, supported by a number of comments left by participants, who justified their neuter assignment by stating that languages are always neuter (i.e. regardless of their form), e.g. *“Sprachennamen sind immer sächlich, z.B. das Deutsch, das Englisch, das Spanisch”*.

Another possible explanation for the relatively high number of participants selecting feminine for *Vanade* is that given by Carstensen (1980: 20) (outlined above), due to the feminine gender of the pseudo-noun’s hypernym, *Sprache* ‘language’, as specified in the definition.

Syllenage (metal): 2(M), 57(F), 34(N), 2(M/N), 5(F/N)

As with *Vanade*, the results for *Syllenage* are largely distributed between feminine and neuter. This is an expected consequence of the presence of the two constraints $-\text{AGE} \rightarrow *M, *N$ and $\text{METALS} \rightarrow *M, *F$. The five participants selecting both feminine and neuter further demonstrates the conflict between the pseudo-noun’s meaning and form. One of these five commented *“das Wort würde ich eindeutig weiblich einschätzen, mir fallen aber nur sächliche Wörter für Metalle ein [I would definitely guess the word to be feminine, but I can only think of neuter words for metals]”*, thereby elucidating the conflict.

The number of participants selecting feminine for *Syllenage* is higher than the number opting for neuter, suggesting a crucial equal ranking between the two constraints $-\text{AGE} \rightarrow *M, *N$ and $\text{METALS} \rightarrow *M, *F$ (or alternatively a ranking whereby $-\text{AGE} \rightarrow *M, *N$ dominates $\text{METALS} \rightarrow *M, *F$). That is to say, the result does not support the hypothesis that $\text{METALS} \rightarrow *M, *F$ is ranked alongside the constraints determining the gender of nouns with referential gender and thus above $-\text{AGE} \rightarrow *M, *N$. However, the difference between feminine and neuter is only weakly statistically significant ($p < 0.05$), meaning that further corroborating evidence would be required before a firm conclusion could be drawn.

Aranne (river): 5(M), 92(F), 1(N), 2(M/F)

A significantly high proportion of participants selected feminine for this pseudo-noun, supporting the proposals that, firstly, $-NNE [N(\partial)] \rightarrow *M, *N$ exists as a constraint, and secondly, that it is ranked above $RIVERS\ NON\text{-}GER. \rightarrow *F, *N$. This ranking is illustrated by a comment left by one participant: “*eigentlich mask. weil Fluss, aber die -nne Endung triumphierte* [actually masculine because it’s a river, but the *-nne* ending triumphed]” (response: F). The second most popular response (although comparatively insignificant) was masculine, with two participants selecting both masculine and feminine in their answer, which demonstrates the form/meaning conflict.

Zabe (fabric): 27(M), 49(F), 15(N), 1(M/N), 4(F/N), 3(M/F), 1(M/F/N)

The majority of participants assigned feminine to *Zabe*, and the number of those who did so is moderately significantly higher than the number of those assigning *Zabe* masculine, the second most commonly selected gender ($p < 0.01$). This provides relatively strong evidence to support the claim that $-E \rightarrow *M, *N$ is ranked above $FABRIC \rightarrow *F, *N$, as was suggested in 3.4.2 on the basis of the noun in the sample *die Seide*. This is further supported by the feminine gender of other fabrics in the German lexicon with a final schwa such as *die Viskose*, *die Baumwolle* and *die Chenille*. An alternative explanation for the result achieved, however, could be that $FABRIC \rightarrow *F, *N$ is not a psychologically real constraint for German native speakers, although this seems unlikely given the fact that the second most commonly selected gender was masculine, which would suggest a conflict between meaning and form.

Of all the pseudo-nouns tested, the distribution of the results obtained for *Zabe* was the most even across the three genders (although statistically still highly significantly different from a 33:33:33 distribution). This is in part due to a moderate number of results for neuter, which might perhaps be explained using Carstensen’s (1980: 20) hypernym theory, based on the hypernym *das Gewebe*. While the hypernym provided in the definition was in fact *der Stoff*, participants may have been reminded of *Gewebe* due to the final *-be* of *Zabe* (although this was not indicated by the pre-experiment test, see 4.3).

Bazka (currency): 51(M), 36(F), 3(N), 9(M/F), 1(M/F/N)

For *Bazka*, the results were almost entirely distributed between masculine and feminine, which is expected due to the two relevant constraints $MON.\ UNITS \rightarrow *F, *N$ and $-A \rightarrow *M, *N$. Nine participants selected both masculine and feminine, further highlighting the conflict between the meaning and form of this pseudo-noun. While the number of participants opting for masculine is slightly higher than the number opting for feminine, the difference between the two results is statistically insignificant

($p < 0.2$). The p -value for *Bazka* was in fact the highest of all pseudo-nouns tested. This means that no firm conclusions can be drawn with regard to the hypothesis that $-A \rightarrow *M, *N$ is ranked above $MON. UNITS \rightarrow *F, *N$.

One factor which may well have contributed to the slightly higher result for masculine is the fact that *der Euro* is highly likely to be both the most frequent and the most prototypical currency for German native speakers, which may have led to the reinforcement of the constraint $MON. UNITS \rightarrow *F, *N$, and perhaps for some speakers, the ranking of $MON. UNITS \rightarrow *F, *N$ above $-A \rightarrow *M, *N$. This was illustrated by a few participants who had selected masculine leaving comments such as “*wie der Euro* [like *der Euro*]”.

Those assigning feminine to *Bazka*, on the other hand, may not only have been influenced by the form of the pseudo-noun, ranking $-A \rightarrow *M, *N$ above $MON. UNITS \rightarrow *F, *N$, but also the feminine gender of the hypernym *die Währung* ‘currency’ (Carstensen 1980: 20). Additionally, it is perhaps possible that some older speakers may have been influenced by the gender of the formerly highly frequent *die Mark* ‘Deutschmark’, thereby instead perhaps applying the constraint $MON. UNITS \rightarrow *M, *N$ to *Bazka*.

***Schnucht*: 10(M), 79(F), 5(N), 1(F/N), 5(M/F)**

A significant majority of participants selected feminine for *Schnucht*, providing evidence to confirm the hypothesis that $-FT/-CHT \rightarrow *M$ is ranked above $[[C]-MONO \rightarrow *F, *N$ (and $CCVCC \rightarrow *F, *N$ if it exists, see *Placht*). The second most commonly selected answer was masculine (albeit insignificantly so), and five participants stated that both masculine and feminine were plausible genders for *Schnucht*, thereby illustrating the competition between the two constraints $-FT/-CHT \rightarrow *M$ and $[[C]-MONO \rightarrow *F, *N$, which arises from a conflict between gender assignment based on the onset and gender assignment based on the coda. The selection of feminine by the majority of participants is therefore an instance of the dominance of the coda over the onset, which is consistent both with Köpcke’s (1982) proposed hierarchy of phonological gender assignment principles for monosyllabic nouns (CODA » STAND-BY » ONSET » STRUCTURE » NUCLEUS) (see 1.3.2 and 1.4.1) and also with the broader claim that the end of a word is more perceptually salient than its beginning (e.g. Slobin 1973).

***Placht*: 13(M), 73(F), 4(N), 1(M/N), 1(F/N), 5(M/F), 3(M/F/N)**

The significantly most commonly selected gender for *Placht* was feminine, supporting the hypothesis that either $-FT/-CHT \rightarrow *M$ is ranked above $CCVCC \rightarrow *F, *N$, or that $CCVCC \rightarrow *F, *N$ is not a psychologically real principle for native speakers, as suggested in 3.4.2. The second most commonly selected gender (although insignificantly so) was masculine, and five participants selected both feminine and

masculine as possible answers. This indicates a possible conflict between the two constraints -FT/-CHT→*M and CCVCC→*F,*N, or a conflict between the constraint -FT/-CHT→*M and masculine as the system default. If CCVCC→*F,*N were a psychologically real constraint, then this would be an instance of gender assignment based on the coda dominating gender assignment based on the syllable structure, which is consistent with Köpcke's (1982) aforementioned proposed hierarchy of phonological gender assignment principles for monosyllabic nouns (see 1.4.1).

***Grett*: 17(M), 7(F), 57(N), 13(M/N), 1(F/N), 2(M/F), 3(M/F/N)**

The highest result obtained for *Grett* was for neuter, which was statistically very significantly higher than the result obtained for the second most commonly selected gender, masculine ($p < 0.001$). This indicates that either -ETT→*M,*F is ranked above CCVC MONO→*F,*N, or that CCVC MONO→*F,*N is not a psychologically real constraint, as suggested in 3.4.2. The reasonably high number of participants selecting both masculine and neuter as possible genders for *Grett* could be explained either by the conflict between the two constraints, or by the conflict between -ETT→*M,*F and masculine as the system default. If CCVC→*F,*N were a psychologically real constraint, then this would be an instance of gender assignment based on a “stand-by” feature (-ETT→*M,*F) dominating gender assignment based on the syllable structure, which is consistent with Köpcke's (1982) aforementioned proposed hierarchy of phonological gender assignment principles for monosyllabic nouns (see 1.4.1).

***Jaff (fruit)*: 28(M), 51(F), 11(N), 4(M/N), 2(F/N), 1(M/F), 3(M/F/N)**

The highest result obtained for *Jaff* was feminine, which was moderately significantly higher than the result obtained for masculine, the second most commonly selected gender ($p < 0.02$). This supports the claim that either FRUIT→*M,*N is ranked above CVC MONO→*F or that CVC MONO→*F is not a psychologically real constraint. The moderately high proportion of results for masculine could either be a result of competition between the two constraints FRUIT→*M,*N and CVC MONO→*F (i.e. meaning and form), or competition between FRUIT→*M,*N and masculine as the system default.

Relatively speaking, *Jaff* had one of the more even distributions of results across the three genders, (although statistically still highly significantly different from a 33:33:33 distribution). This is owed in part to the moderate number of results for neuter, which might perhaps be explained by Carstensen's (1980: 20) hypernym theory, based on the neuter gender of *das Obst* ‘fruit’.

Felpt (game): 13(M), 1(F), 78(N), 7(M/N), 1(F/N)

A significant majority of participants assigned neuter to *Felpt*, thereby supporting the claim that CVCCC MONO→*F,*N is not a psychologically real constraint for native speakers, as suggested in 3.4.2. An alternative explanation for the result could, however, be that the constraint GAMES→*M,*F is ranked above CVCCC MONO→*F,*N. The relatively small number of participants selecting masculine or selecting both masculine and neuter as possible genders might either indicate a tension between the two constraints (i.e. meaning and form), or a tension between GAMES→*M,*F and masculine as the system default gender.

Henn (unit): 19(M), 9(F), 67(N), 2(M/N), 2(F/N), 1(M/F)

A significant majority of participants assigned *Henn* neuter, providing evidence to support the proposal that -N(C)(C) MONO→*F,*N is not a psychologically real constraint for native speakers. An alternative explanation for the results could, however, be that the constraint SCI. UNITS→*M,*F is ranked above -N(C)(C) MONO→*F,*N. The fact that the second most common response was masculine indicates a conflict either between SCI. UNITS→*M,*F and -N(C)(C) MONO→*F,*N (i.e. meaning and form), or between SCI. UNITS→*M,*F and masculine as the system default.

One comment left by a participant who selected masculine was also illuminating: “*mask. als Maßeinheit [masculine as a unit of measurement]*”, suggesting that at least some speakers might instead consider scientific units to be associated with masculine, i.e. apply the constraint SCI. UNITS→*F,*N. This could be explained by the fact that some highly prototypical and frequent units are often used with masculine targets, e.g. *Meter* and *Liter* (Duden Grammatik 2009: 234, 236). Moreover, in order to acquire the constraint SCI. UNITS→*M,*F, it is clear that a certain level of scientific knowledge is required, given that most units (e.g. *das Hertz, das Joule, das Newton, das Farad*) occur at most infrequently in everyday language. Indeed, some of the participants who assigned neuter to *Henn* added comments such as: “*wie das Becquerel, das Newton, das Pascal etc. [like das Becquerel, das Newton, das Pascal etc.]*”, thereby demonstrating scientific knowledge. Those without this knowledge therefore may not have the constraint SCI. UNITS→*M,*F in their gender assignment system, and may have assigned *Henn* masculine by default.

Finally, the (relatively small) number of responses for feminine might be explained by the fact that *die Einheit* ‘unit’ was the hypernym given in the definition (see Carstensen 1980: 20).

***Lasp* (insect): 31(M), 49(F), 12(N), 2(F/N), 5(M/F), 1(M/F/N)**

Most participants assigned feminine to *Lasp*, the number doing so being statistically significantly higher than those selecting masculine (the second most common response), but only weakly so ($p < 0.05$). This therefore constitutes evidence in favour of the claim that CVCC MONO→*F is not a psychologically real constraint. Alternatively, the result could be explained by the constraint INSECTS→*M,*N being ranked above CVCC MONO→*F. Masculine as the second most common response and the fact that five participants considered both masculine and feminine to be plausible genders for *Lasp* indicates either a tension between the two constraints (i.e. between meaning and form), or between INSECTS→*M,*N and masculine as the system default.

Lasp was among the pseudo-nouns with a relatively more even distribution of results across the three genders (although ultimately highly significantly different from a completely even distribution). This could partly be explained by the moderate number of responses for neuter, which may be due to Carstensen's (1980: 20) theory, given that the hypernym provided in the definition was *das Insekt*.

***Ahb* (alcoholic drink): 56(M), 0(F), 32(N), 7(M/N), 3(M/F), 2(M/F/N)**

The results for *Ahb* were almost entirely distributed between masculine and neuter, showing the effects of the constraints ALC. DRINKS→*F,*N, V_{LONG}(C) MONO→*F and VC MONO→*M,*F. The number opting for masculine was moderately significantly larger than the number opting for neuter ($p < 0.01$), which provides relatively strong evidence to support the proposal that the three constraints are crucially equally ranked. An alternative explanation for the results obtained would be that ALC. DRINKS→*F,*N were ranked above VC MONO→*M,*F, however, there is no independent reason to suggest this.

The relatively high proportion of neuter results and the seven results selecting both masculine and neuter illustrates the conflict between the two constraints ALC. DRINKS→*F,*N and VC MONO→*M,*F, i.e. the conflict between the pseudo-noun's meaning and form. The neuter results might also be explained in part by Carstensen's (1980: 20) hypothesis, since the hypernym given for *Ahb* was *das Getränk* 'drink'.

***Fralang* (dwelling): 60(M), 4(F), 26(N), 8(M/N), 1(F/N), 1(M/F/N)**

The gender most commonly assigned to *Fralang* was masculine, the number doing so being greater to a high level of significance ($p < 0.001$) than the number selecting neuter, which was the second most common response. This result is consistent with the existence of the two constraints -ANG→*F,*N and DWELL→*M, as well as a *N»*F»*M default hierarchy. Possible alternative explanations for the result

achieved could be that DWELL→*M is not a psychologically real constraint and *Fralang* was assigned masculine purely due to -ANG→*F,*N, or that neither DWELL→*M nor -ANG→*F,*N are psychologically real constraints, and *Fralang* was assigned masculine by default, although there is no independent reason to suggest that either of these are true.

The number of participants selecting the second most common response – neuter, or selecting both masculine and neuter as possible genders for *Fralang* might be accounted for in a number of ways. One possible explanation is that the constraint DWELL→*M is in fact the originally proposed constraint DWELL→*M,*F (see Appendix A), which was modified to DWELL→*M in 2.9 due to failing the first criterion (with just a 48% coverage of candidate nouns). These results then might be accounted for by a conflict between the constraints DWELL→*M,*F and -ANG→*F,*N. This could be due to the prototypicality and the high frequency of *das Haus*, which may have led to the reinforcement of the constraint DWELL→*M,*F. Another possible explanation for the neuter results is Carstensen’s (1980: 20) hypernym theory, given the neuter hypernym *das Haus*. Nevertheless, such hypotheses do not have sufficient statistical support, given that the number of participants selecting neuter is insignificant in comparison to the number of masculine responses.

***Klepisma*: 5(M), 17(F), 68(N), 1(M/N), 5(F/N), 1(M/F), 3(M/F/N)**

A statistically significant majority of participants selected neuter for *Klepisma*, which supports the proposals that, firstly, -ISMA→*M,*F is a psychologically real constraint in the gender assignment system of native German speakers, and, secondly, that -ISMA→*M,*F is ranked above -A→*M,*N. The fact that the second highest response was feminine, and that five participants selected both feminine and neuter for their answer, demonstrates the conflict between these two constraints.

***Grileau*: 28(M), 6(F), 54(N), 8(M/N), 1(F/N), 3 (M/F/N)**

The number of participants assigning *Grileau* neuter is statistically moderately significantly higher than the number of those who selected masculine, the second most common response ($p < 0.01$). This constitutes relatively strong evidence for the existence of the constraint -EAU→*M,*F.

The proportion of participants selecting masculine might be due to the pseudo-noun being perceived as a French noun and being assigned masculine on the basis that French nouns in *-eau* are usually masculine (Rosenthal 2009: 25). This is evidenced by a number of comments left by participants, such as: “*franz. männliche Endung* [French masculine ending]”, “*-eau = m. Französisch* [-*eau* = m. French]”. There was, however, some explicit recognition of *-eau* being associated with masculine in French but

neuter in German, with one participant who selected neuter stating: “*Französisch: -eau = mask. aber Deutsch: -eau = neut.* [French: *-eau = masculine but German -eau = neuter*]”.

Another possible explanation for the non-neuter responses is that the number of nouns ending in *-eau* in German is relatively small – there are just 14 such nouns listed in the Duden-Bibliothek Express. This means that the constraint $-EAU \rightarrow *M, *F$ is unlikely to be reinforced very often.

A final plausible explanation for the non-neuter responses arises as the result of one of the major limitations of the study. The fact that the pseudo-nouns were presented in a written format rather than in an audible form means that no stress patterns could be indicated. As discussed in 1.3.2, stress patterns can be highly relevant to gender assignment, e.g. STRESSED $-ON \rightarrow *F, *N$ as in *der Balkon* vs. UNSTRESSED $-ON \rightarrow *M, *F$ as in *das Stadion* (Nelson 1998). In the case of the constraint $-EAU \rightarrow *M, *F$, stress is indeed relevant as the constraint is only applicable to nouns with a stressed final *-eau* [‘o:]. Should the pseudo-noun have been analysed as having stress on the initial syllable, the constraint $-EAU \rightarrow *M, *F$ would not have been applied by the participant, resulting in non-neuter gender assignment.

***Trelair*: 20(M), 12(F), 55(N), 7(M/N), 2(F/N), 1(M/F), 3(M/F/N)**

The majority of participants assigned neuter to *Trelair*, the number of those who did so being greater to a high level of statistical significance than the number who assigned masculine to the pseudo-noun, which was the second most common response ($p < 0.001$). This therefore constitutes strong evidence for the existence of the constraint $-AIR [\varepsilon:\text{ɛ}] \rightarrow *M, *F$.

As with *Grileau*, the number of participants selecting masculine or selecting both masculine and neuter as possible genders for *Trelair* is likely to be due to the perception of the pseudo-noun as French, since French nouns in *-air* are usually masculine (Rosenthal 2009: 40). The non-neuter responses might also be explained by the relatively uncommon occurrence of final *-air* in German (with just 11 nouns of this type in the Duden-Bibliothek Express), leading to infrequent reinforcement of the constraint $-AIR [\varepsilon:\text{ɛ}] \rightarrow *M, *F$. Finally, the lack of the provision of an audible form of the pseudo-noun may also have contributed to non-neuter assignment, since the constraint $-AIR [\varepsilon:\text{ɛ}] \rightarrow *M, *F$ would not have applied if *Trelair* had been perceived as having stress on the initial syllable, or if the final segment had not been perceived as $[\varepsilon:\text{ɛ}]$, but instead as e.g. $[a\text{?i}:\text{ɪ}]$.

***Dranalt*: 69(M), 8(F), 11(N), 4(M/N), 5(M/F), 3(M/F/N)**

The majority of participants assigned masculine to *Dranalt*, the number doing so being greater to a high level of statistical significance than the number selecting neuter, which was the second most common response ($p < 0.001$). The results obtained are therefore consistent with the existence of

both the constraint $-ALT \rightarrow *N$ and a default markedness hierarchy in which masculine is considered to be less marked than feminine.

The same results could have been obtained with a default hierarchy consisting of $*F \gg *M \gg *N$ rather than $*N \gg *F \gg *M$ as is proposed to be the case, however this is in general a less plausible ranking, as demonstrated in 3.4.3. The second most common response was neuter, although the number doing so is not statistically significantly different from the number selecting feminine ($p < 0.5$). With five participants selecting both masculine and feminine as possible genders for *Dranalt*, there is some evidence of a conflict between the effects of the constraint $-ALT \rightarrow *N$ and those of the markedness hierarchy.

Region as a variable

As stated in 4.3, all participants were asked to specify their region of origin before completing the questionnaire. The participants originated from all areas of Germany, with at least one participant from each federal state except for Saarland. The majority of participants originated from Bavaria, Baden-Württemberg and North Rhine-Westphalia, the three most populous states in Germany.

While region of origin is not the main variable being examined in this study, it was nonetheless considered relevant to obtain this information, so that region could be investigated as a possible factor for certain cases of gender assignment, e.g. the South-West and West-Central German tendency to use neuter targets with women's names (see discussion of *Yall*).

In order to determine whether region of origin may have had a broader effect on the distribution of results obtained for each pseudo-noun, a statistical test was performed. Firstly, the participants' results were divided into six large dialect areas (see e.g. Wiesinger 1983, König 2001): West Low German (19 participants), East Low German (6 participants), West-Central German (12 participants), East-Central German (11 participants), Alemannic (17 participants) and Bavarian (35 participants). These are the same dialect areas used in e.g. Callies et al.'s (2012) study of regional variation in gender assignment to Anglicisms.

The distribution of results obtained for each pseudo-noun was then calculated for each region, and subsequently compared to the overall distribution of results for each pseudo-noun using a chi-square test. It was found that the overall distribution of results was not statistically significantly different from the distribution of results obtained for any of the pseudo-nouns in any of the six dialect areas ($p > 0.05$). That is to say, the distribution of results obtained for each pseudo-noun was similar in each region.

However, it is important to note that in order for any firm conclusions to be drawn about regional variation in gender assignment to pseudo-nouns, a larger and more balanced sample would be required (cf. Callies et al. (2012), who used a sample of 506 native speakers, with approximately the same number of speakers from each area being investigated).

Other sociolinguistic factors

While further sociodemographic information about the participants was not collected and thus cannot be investigated for its possible effect on gender assignment, the results suggest that at least some sociodemographic factors may have influenced the genders selected by the participants. One such factor is the level of education a participant has received. For instance, for both *Grileau* and *Trelair*, it was suggested (and illustrated by a number of comments) that metalinguistic knowledge of French may have contributed to the number of participants selecting masculine for these nouns. Similarly, for *Henn*, it was proposed that the application of the constraint SCI. UNITS→*M,*F may only be possible for those with a certain degree of scientific knowledge.

This notion is consistent with claims made in a number of studies on gender assignment to Anglicisms, e.g. Schlick (1984: 421), Schulte-Beckhausen (2002: 77), Fischer (2005: 279), who maintain that an individual's educational level affects their selection of noun gender. It would therefore be interesting to explore this further.

4.6 Conclusions and implications of results

In summary, the null hypothesis can be decisively rejected on the basis of the data above. That is to say, it has been shown that the semantic, morphological and phonological features of the pseudo-nouns did indeed influence the gender assigned to them by native German speakers. This is evident because for all pseudo-nouns tested, the difference between an entirely random distribution of results across the three genders and the actual distribution of results was highly statistically significant ($p < 0.001$).

In almost all cases, the exact way in which the pseudo-nouns' features influenced the gender assigned to them was as expected, meaning that the results support the majority of the specific proposals which were outlined in 4.2. The only pseudo-nouns for which this was not the case were *Syllenage*, *Vanade* and *Bazka*. For *Syllenage*, the results suggest a crucial equal ranking between $-AGE \rightarrow *M,*N$ and $METALS \rightarrow *M,*F$ instead of a state of affairs in which $METALS \rightarrow *M,*F$ is ranked alongside the constraints determining the gender of nouns with referential gender and thus above $-AGE \rightarrow *M,*N$. For *Vanade*,

although the results were consistent with the hypothesis that LANGUAGES→*M,*F is ranked above -ADE→*M,*N, they were not sufficiently statistically significant to confirm it. Similarly, for *Bazka*, the results were not sufficiently statistically significant to draw any firm conclusions about the ranking of -A→*M,*N and MON. UNITS→*F,*N.

The results otherwise provide evidence to support the hypotheses stated in 4.2. That is to say, the results confirm that many of the adaptations to OGAT proposed in Chapter 3 are necessary in order to account for the data, i.e. that OGAT II, a model in which certain GENDER FEATURES constraints are ranked higher than others, although marginally less elegant than the original version of OGAT, is empirically justified.

In general, the results show relatively high levels of agreement amongst native speakers as to the genders of the pseudo-nouns, which further supports the argument that all native speakers possess a gender assignment system which enables them to “work out” the gender of a noun based on its features (e.g. Corbett 1991: 7) (see 1.2).

The results also provide evidence that the pseudo-nouns investigated were not simply assigned gender on the basis of analogy with similar-sounding nouns, as is a concern raised by e.g. Köpcke & Zubin (1983), Mills (1986), Wegener (1995), Hohlfeld (2006) etc. This is apparent because, firstly, there are multiple cases of the gender most commonly assigned being distinct from the genders of the similar-sounding nouns listed by native speakers in the pre-experiment test, and secondly, there are many cases in which the form of the pseudo-noun does not appear to be the decisive factor in determining its gender.

That said, for many of the pseudo-nouns, there is at least some evidence to suggest that both the meaning and the form of the pseudo-noun were taken into consideration by the participants when assigning gender, thus supporting the claim that a native speaker’s gender assignment system uses both meaning and form-based regularities (e.g. Zubin & Köpcke 1981, Köpcke 1982, Corbett 1991) (see 1.3). Where the results are largely distributed between two genders, or where quite a number of participants selected two genders in their answer, one of these genders is usually the gender associated with the pseudo-noun’s form, and the other is that which is associated with the pseudo-noun’s meaning.

Furthermore, the results provide additional evidence against the claim that gender assignment principles are ranked according to type (see 1.4 and 3.5.3). For instance, the results for the pseudo-nouns *Yällchen*, *Springling*, *Randor*, *Mellade*, *Syllenage*, *Aranne*, *Zabe* and *Fralang* constitute evidence against the proposal that semantic principles always override form-based principles (e.g. Corbett

1991), since it was the form of these nouns which ultimately determined their gender instead of their meaning. In addition, the results for *Zunie, Yall, Gamie, Arlt, Jaff, Felpt, Henn, Lasp* and *Ahb* constitute evidence against the reverse proposal, that form-based principles always override semantic principles, since it was the meaning, and not the form, of these nouns which ultimately determined their gender. Finally, the results obtained for *Mellade, Syllenage, Aranne, Zabe* and *Fralang* constitute evidence against the claim that morphological principles override semantic principles which in turn override phonological principles (e.g. Wegener 1995), since the genders of these nouns were ultimately determined by their phonological form rather than their meaning. This further demonstrates that the OGAT approach to gender assignment is a better representation of the German system than approaches in which constraints are ranked according to their broad type.

Additionally, the results obtained for *Schnucht, Placht* and *Grett* appear to support Köpcke's (1982) proposed hierarchy of phonological gender assignment principles for monosyllabic nouns (CODA » STAND-BY » ONSET » STRUCTURE » NUCLEUS, see 1.4.1), which is therefore something that could be incorporated into OGAT II subject to further testing.

In terms of the specific implications of the results, we have seen that the results support the ranking of $-CHEN \rightarrow *M, *F$ above a block of constraints (BLOCK 2) which determine the gender of nouns with referential gender, such as $NAME:WOMAN \rightarrow *M, *N$ and $CITIES/TOWNS \rightarrow *M, *F$. The results also suggest that $LETTERS/MUS. NOTES \rightarrow *M, *F$; $CHEESES \rightarrow *F, *N$; and (possibly) $LANGUAGES \rightarrow *M, *F$ belong in BLOCK 2 as well.

The results then support the ranking of BLOCK 2 above BLOCK 3, a block of "strong" constraints which – apart from in cases of referential gender – otherwise usually categorically determine the gender of a noun, such as $-IE [i:] / [i\text{a}] \rightarrow *M, *N$; $-LING \rightarrow *F, *N$; $-OR \rightarrow *F, *N$; $-AGE \rightarrow *M, *N$; $-ADE \rightarrow *M, *N$; and $VCCC MONO \rightarrow *M$. The results also suggest that $METALS \rightarrow *M, *F$ is crucially equally ranked alongside these constraints in BLOCK 3.

Lastly, the results confirm that ranked below BLOCK 3 is a further block of constraints, BLOCK 4, which contains the remaining GENDER FEATURES constraints tested in the study, including $ALC. DRINKS \rightarrow *F, *N$; $-L(C) MONO \rightarrow *F$; $-ANG \rightarrow *F, *N$; $FLOWERS \rightarrow *M, *N$; $TREES \rightarrow *M, *N$; $VC MONO \rightarrow *M, *F$ and $FISH \rightarrow *F, *N$. The results also provide evidence for the existence of the following constraints in BLOCK 4: $-EAU \rightarrow *M, *F$; $-AIR [\text{ɛ}:\text{ɛ}] \rightarrow *M, *F$; $-NNE [N(\text{ə})] \rightarrow *M, *N$ and $-ALT \rightarrow *N$.

The results support the claim that all constraints are crucially equally ranked with respect to one another within BLOCK 4, with a small number of exceptions. Specifically, the results endorse the

following rankings: $-ISMA \rightarrow *M, *F \gg -A \rightarrow *M, *N$; $-E \rightarrow *M, *N \gg FABRIC \rightarrow *F, *N$; $-NNE [N(\emptyset)] \rightarrow *M, *N \gg RIVERS$
 $NON-GER. \rightarrow *F, *N$; and $-FT/-CHT \rightarrow *M \gg [j]C- MONO \rightarrow *F, *N$.

There are a number of constraints which, in line with the results, either are not psychologically real constraints for native speakers, or, if they do exist, are ranked below particular constraints in BLOCK 4. These are $CCVCC \rightarrow *F, *N$; $CCVC MONO \rightarrow *F, *N$; $CVC MONO \rightarrow *F$; $CVCCC MONO \rightarrow *F, *N$; $-N(C)(C) MONO \rightarrow *F, *N$; and $CVCC MONO \rightarrow *F$.

Finally, the results support the default markedness hierarchy $*N \gg *F \gg *M$, which is ranked below all GENDER FEATURES constraints.

A summary of what the results support can be seen in tableau format in Figure 4 overleaf.

Input	GENDER FEATURES				DEFAULT HIERARCHY		
	BLOCK 1 (DIMINUTIVE)	BLOCK 2 ("REFERENTIAL")	BLOCK 3 ("STRONG")	BLOCK 4 (OTHER)			
	-CHEN→*M,*F	CITIES/TOWNS→*M,*F; NAME:WOMAN→*M,*N; LETTERS/MUS. NOTES→*M,*F; CHEESES→*F,*N; (LANGUAGES→*M,*F)	-ADE/AGE→*M,*N; -IE ['i:]/[iə]→*M,*N; -LING→*F,*N; METALS→*M,*F; -OR→*F,*N; VCCC MONO→*M	ALC. DRINKS→*F,*N; DWELL→*M; FLOWERS→*M,*N; TREES→*M,*N; FISH→*F,*N; MON. UNITS→*F,*N; -L(C) MONO→*F; VC MONO→*M,*F; V _{LONG} (C) MONO→*F; -A→*M,*N; -AIR [ɛ:ə]→*M,*F; -ANG→*F,*N; -ALT→*N; -EAU→*M,*F; -ISMA→*M,*F » -A→*M,*N; -E→*M,*N » FABRIC→*F,*N; -NNE [N(ə)]→*M,*N » RIVERS NON-GER.→*F,*N; -FT/CHT→*M » []C-MONO→*F,*N;(CCVCC→*F,*N); -ETT→*M,*F (» CCVC MONO→*F,*N); FRUIT→*M,*N (» CVC MONO→*F); GAMES→*M,*F (» CVCC MONO→*F,*N); SCI. UNITS→*M,*F (» -N(C)(C) MONO→*F,*N); INSECTS→*M,*N (» CVCC MONO→*F)	*N	*F	*M
<i>der</i> ____							*
<i>die</i> ____						*	
<i>das</i> ____					*		

Figure 4: Summary of what the results support in tableau format

Since the results largely confirm the hypotheses outlined in 4.2, no significant changes need to be made to OGAT II on the basis of the results obtained. The small adaptations that are required, e.g. the movement of METALS→*M,*F to BLOCK 3, do not have any negative impact on OGAT II's ability to account for the genders of 94.9% of the nouns in the sample.

In fact, the results support a number of changes which have the potential both to improve OGAT II's ability to account for the genders of the nouns in the sample, as well as to simplify the model somewhat. Specifically, these changes involve the elimination of the six constraints CCVC→*F,*N; CCVC MONO→*F,*N; CVC MONO→*F; CVCC MONO→*F,*N; -N(C)(C) MONO→*F,*N; and CVCC MONO→*F. Should these constraints be eliminated from the model (rather than appear ranked below particular constraints, which is the alternative explanation for the results obtained), then OGAT II would be able to account for the genders of two additional nouns in the sample: *Bank* and *Macht*. Overall, this would mean that OGAT II would be able to account for the genders of 95.3% of the nouns in the sample.

Furthermore, the model would be simplified, since not only would it require fewer constraints, but it would also require eight fewer ranking arguments in BLOCK 4, which were originally proposed in 3.3.4.4 and 3.4.2 in order to account for nouns such as *Flett*, *Steelband*, *Kunst*, *Pimperness*, *Benediktenkraut* and *Schwestertracht*. These are: (1) -ETT→*M,*F » CCVC MONO→*F,*N; (2) -FT/-CHT→*M » CCVC→*F,*N; (3) FRUIT→*M,*N » CVC MONO→*F; (4) SUP.→*M,*F » CCVC MONO→*F,*N; (5) ACAD. DISCIPLINES→*M,*N » -N(C)(C) MONO→*F,*N; (6) ACAD. DISCIPLINES→*M,*N » CVCC MONO→*F,*N; (7) GROUPS→*M,*N » CVCC MONO→*F; and (8) GROUPS→*M,*N » -N(C)(C) MONO→*F,*N. Without the six above-listed constraints, none of these rankings would be necessary.

The final model, which incorporates all of the changes that have been suggested on the basis of the results of this experiment, is presented in Figure 5.

Input	GENDER FEATURES					DEFAULT HIERARCHY			
	BLOCK 1 (DIMINUTIVE)	BLOCK 2 ("REFERENTIAL")	BLOCK 3 ("STRONG")	BLOCK 4 (OTHER)		BLOCK 5 (SEM. EQUIV.)	*N	*F	*M
-CHEN→*M,*F	NAME:MAN→*F,*N; NAME:WOMAN→*M,*N; CITIES/TOWNS→*M,*F; LANGUAGES→*M,*F; LETTERS/MUS. NOTES→*M,*F; CHEESES→*F,*N	METALS→*M,*F -(ER)EI→*M,*N -(O)LOGE→*F,*N -ADE/AGE→*M,*N -ANER→*F,*N -ANZ/ENZ→*M,*N CONVERSION→*M,*F -EDER→*M,*F -EKT→*F -ER/OR SUFF.→*F,*N -HEIT/- (IG)KEIT→*M,*N -IE ['i:]/[iə]→*M,*N -IN SUFF.→*M,*N -INER→*F,*N -ION→*M,*N -IP [i:p]→*M,*F -ISMUS→*F,*N -IST→*F,*N -ITIS→*M,*N -IUM→*M,*F -LING→*F,*N -MENT→*M,*F -NIS→*M -ON UNSTRESS.→*M,*F -SAL→*M -SCHAFT→*M,*N -SKOP→*M,*F -TÄT→*M,*N -THEK→*M,*N -UNG→*M,*N -UR SUFF.→*M,*N VCCC MONO→*M	-(C)(C)/[/] MONO→*F,*N -/r/ STOP (C) MONO→*F,*N [[]C- MONO→*F,*N -A→*M,*N ACAD. DISCIPLINES→*M,*N -AIR [ɛ:ə]→*M,*F -AL→*F ALC. DRINKS→*F,*N -AL ENG.→*M,*F -ALT→*N -AN→*M,*F -ANG→*F,*N -ANT/-ENT→*F,*N -AR→*M,*F -AT→*M,*F -B→*F BIRDS→*F -C S MONO→*F,*N CCVC MONO→*F,*N CCVV _{DIPHTHONG} - MONO→*F COUNTRIES→*M,*F CV MONO→*N DAYS→*F,*N DEVERBAL STEM→*F,*N DISEASES→*M/→*N DOM. ANIMALS→*F DR→*F,*N DWELL→*M -E→*M,*N -EAU→*M,*F -EE/-ÉE/-É→*F -EL(L) ['ɛl]→*M,*F (-EN→*F)	-IS→*M,*N -ISMA→*M,*F ISLANDS→*M,*F -IT→*F,*N/-IT→*F -IV→*M,*F -L(C) MONO→*F LEATHER TYPES→*M,*F MALE ANIMALS→*F,*N MAMMALS→*F MEN→*F,*N MON. UNITS→*F,*N MOUNTAINS→*F,*N MUS. TERMS/COMP.→*M MUS. INSTR.→*M MUSIC TYPES→*F,*N -NNE [N(ə)]→*M,*N -O→*F ORGANS→*M/→*N ∅V _{LONG} MONO→*F PICTURES→*M PLANTS→*N PRIMATES→*F,*N R-/T-/D- MONO→*F REGIONS→*M,*F RELIGIONS→*F,*N RIVERS EUROPE→*N RIVERS NON- EUROPE→*F,*N RIVERS NON-GER.→*F,*N ROCKS/MINERALS→*F,*N SCI. UNITS→*M,*F SEASONS→*F,*N -ST/-SCHT/-ZT→*N	SEM. EQUIV. →*M/*F/*N				

Chapter 5: Conclusions and directions for further work

The focus of this project has been the exploration of the notoriously problematic phenomenon of grammatical gender assignment in German. Despite claims that gender assignment is arbitrary and unpredictable in German (e.g. Maratsos 1979, Pfau 2009), the findings of this investigation contribute to the body of evidence which suggests that gender assignment in German is in fact a largely regular, systematic process (e.g. Köpcke 1982, Köpcke & Zubin 1983, 1984, Rice 2006), and more broadly add weight to the claim that gender assignment is 'essentially systematic' in all languages (Corbett 1994: 1350). As well as showing that there are discernible regularities in the German lexicon, the findings of this study also contribute to the evidence which suggests that native speakers productively exploit gender assignment regularities when assigning genders to new lexical items (e.g. Schiller et al. 2003, Schwichtenberg & Schiller 2004).

In theoretical terms, this study has developed and built on Optimal Gender Assignment Theory (OGAT), proposed by Rice (2006) on the basis of earlier work by Steinmetz (1985, 1986, 2001, 2006), and demonstrates that the theory, with a small number of adaptations, is able to account for the genders of a high proportion of German nouns. Specifically, it has been demonstrated that OGAT II, a revised version of OGAT, is able to account for the genders of around 95% of a representative sample of 592 nouns, as well as being able to capture important generalisations about German gender assignment.

A particular strength of this study is that, unlike previous studies of gender assignment, it has established and employed a methodology for testing the validity of GENDER FEATURES constraints (see Chapter 2), thereby overcoming criticisms that earlier studies tended to include ad hoc constraints in order to produce the desired outcomes (see 1.4.5.3). A further strength of the study is that, unlike previous testing of OGAT, the theory has been tested using a substantial noun sample (see 1.4.5.3). The sample was systematically selected and, statistically, almost exactly mirrors the entire Duden corpus in terms of the distribution of nouns across the genders (see 2.2), suggesting that the sample is a good representation of the German lexicon as a whole. Moreover, the testing was conducted on a wide variety of nouns, including common nouns, proper nouns (3.3.4), loanwords (3.3.1), polysemes (3.3.2) and multiple-gender nouns (3.3.4).

The investigation has demonstrated that morphological, phonological and semantic GENDER FEATURES constraints are needed to account for the German data, and has also explored topics such as the direction of derivation to deal with problematic constraints such as DEVERBAL STEM→*F,*N (see 3.2.2.1), and semantic theories to account for issues such as superordinates (see 3.2.2.2) and polysemy (3.3.2). Additionally, the study has shown that OGAT II is most effective with a default hierarchy consisting of

*N»*F»*M, as originally proposed by Rice (2006). Possible alternative rankings were investigated, but these were shown to be less effective than the original hierarchy proposed (see 1.4.5.3 and 3.4.3). This demonstrates that basing the default hierarchy on noun distribution (or, at least, the majority of estimates thereof) accounts best for the German data (see 1.4.5.3).

Furthermore, the results indicate that OGAT II is not only able to account for the genders of a high proportion of existing nouns, but it is also able to make the correct predictions about the genders of novel nouns (see Chapter 4). Specifically, it was found that the predictions made by OGAT II with regard to the genders of a set of 26 pseudo-nouns match the actual results obtained from native German speakers, suggesting that OGAT II accurately models the way in which native speakers assign gender.

Finally, this investigation has demonstrated that OGAT II is better able to account for the German data than many of the other key approaches to gender assignment principle interaction that have been proposed, e.g. the approach in which semantic principles dominate form-based principles (e.g. Köpcke 1982, Corbett 1991), the approach in which form-based principles dominate semantic principles, and the approach in which morphological principles dominate semantic principles which in turn dominate phonological principles (e.g. Heringer 1995, Wegener 1995) (see 1.4 and 3.5.3). This is a highly significant finding, since it contradicts the claims of key gender assignment theorists such as Köpcke (1982) and Corbett (1991).

There is, however, undoubtedly further work to be carried out. The most natural way of building on this study would be to test OGAT II on an even larger sample of nouns, which would also involve the testing of further GENDER FEATURES constraints. Of the 378 gender assignment principles found in the literature, the present study required the examination of 237, meaning that 141 principles remain untested (see 2.2 and Appendix B). However, it is not self-evident that a simple extension of the study would significantly alter the results. That is to say, OGAT II is likely to be able to account for the genders of approximately 95% of nouns even if a larger sample were used. This is because the sample used in the present study has been shown to be largely representative of the lexicon as a whole, having been collected systematically and, moreover, closely mirroring the statistical composition of the entire Duden corpus (see 2.2). Additionally, the results of the present study are based on a sample which contains a wide variety of potentially problematic groups of nouns: loanwords, polysemes, multiple-gender nouns, proper nouns, compounds, affixed nouns, lexicalised noun phrases, archaic nouns, regionalisms etc. Should the sample be expanded, it is thus unlikely that a larger sample would contain many (if any) further problematic large groups of nouns that would significantly affect OGAT II's efficacy.

In terms of the 141 untested GENDER FEATURES constraints (see 2.2 and Appendix B), it is expected that these could be integrated into OGAT II relatively easily, should they emerge as plausible after being tested according to Enger's (2009) six criteria. For example, the constraint $-\text{LEIN} \rightarrow *M, *F$ "nouns ending in the suffix *-lein*, e.g. *das Büchlein*, cannot be masculine or feminine" (e.g. Flämig 1991: 453, Menzel 2004: 66, Hoberg 2004: 87, Weinrich 2007: 326, Engel 2009: 281, Eisenberg 2013: 133) is not relevant to the current sample, but it is expected that this constraint would appear in OGAT II in the diminutive block (block 1, see 4.6) alongside $-\text{CHEN} \rightarrow *M, *F$, since this constraint overrides constraints applicable to nouns with referential gender, e.g. $\text{NAME:WOMAN} \rightarrow *M, *N$, as in *das Annalein* (see 3.3.4.2).

Additionally, it is expected that untested semantic constraints such as $\text{SHIPS} \rightarrow *M, *N$ "names of ships cannot be masculine or neuter, e.g. *die Bismarck*, *die Albatross*" (e.g. Hickey 2000: 630, Hoberg 2004: 108, Köpcke & Zubin 2005a: 119, Eisenberg 2013: 138); $\text{MOTORBIKES} \rightarrow *M, *N$ "nouns denoting makes of motorbike cannot be masculine or neuter, e.g. *die BMW*, *die Honda*" (e.g. Flämig 1991: 452, Menzel 2004: 63, Weinrich 2007: 328, Engel 2009: 272); and $\text{PLANES} \rightarrow *M, *N$ "names of aeroplanes cannot be masculine or neuter, e.g. *die Boeing*, *die Spitfire*" (e.g. Flämig 1991: 451, Hickey 2000: 630, Weinrich 2007: 328, Engel 2009: 272) would appear in OGAT II in the referential block (block 2, see 4.6), since the form of the nouns to which these constraints apply is not analysed; their gender is instead based entirely on their referent (e.g. Nübling 2015).

Other untested constraints such as $-\text{ANCE} [\tilde{a}:s(\text{ə})] \rightarrow *M, *N$ "nouns ending in *-ance* [$\tilde{a}:s(\text{ə})$], e.g. *die Renaissance*, cannot be masculine or feminine" (e.g. Duden Grammatik 2009: 164) and $-\text{ISSE} \rightarrow *M, *N$ "nouns in *-isse*, e.g. *die Kulisse*, cannot be masculine or neuter" (e.g. Weinrich 2007: 327, Duden Grammatik 2009: 165) would appear in OGAT II in the "strong" constraints block (block 3), since all nouns with these features (except for nouns with referential gender) are assigned feminine (based on samples taken from the Duden-Bibliothek Express).

Many of the remaining untested constraints, such as $\text{GR- MONO} \rightarrow *F, *N$ "monosyllables with the onset /gr/ cannot be feminine or neuter, e.g. *der Griff*" (Köpcke 1982: 91) and $\text{KN- MONO} \rightarrow *F, *N$ "monosyllables with the onset /kn/ cannot be feminine or neuter, e.g. *der Knast*, *der Knopf*" (e.g. Köpcke 1982: 89, Mills 1986: 33, Menzel 2004: 68, Hoberg 2004: 97) would appear in OGAT II in the "other constraints" block (block 4), since these constraints do not have 100% coverage of candidate nouns.

As shown above, none of the untested constraints is of a different order from those which have been tested in the present study, and it is therefore unlikely that the untested constraints will cause any significant problems for OGAT II.

A further way to extend the study would be to investigate OGAT II's ability to account for gender in non-standard varieties of German or indeed in other languages. The present study is based on Standard German data alone, and only takes regional variants into account where these are specified in the Duden (see 3.3.4.4 and 3.3.4.8). It would therefore be interesting to establish whether OGAT II is able to account for further regionalisms. Additionally, it would be interesting if the results achieved for Standard German could be replicated for other languages. Preliminary analyses by Rice (2004, 2006) indicate that the OGAT approach to gender assignment can also successfully account for Russian, French, Dutch and Norwegian data. However, these findings are based on small data sets. Further investigation of these languages using larger samples is therefore necessary.

As stated in 1.2, an additional area of further research is the psycholinguistic investigation of the extent to which native speakers exploit gender assignment regularities in language production. The results from the present study indicate that gender assignment regularities – specifically in the form of OGAT II – are used by native speakers when assigning gender to pseudo-nouns. However, only further empirical testing is able to conclusively determine the extent to which the system is employed by native speakers in their assignment of gender to existing nouns (e.g. Corbett 1991, van Berkum 1996, Conzett 2006). It would also be interesting to explore gender assignment outside of the realm of rule/constraint-based approaches examined in the present study. Further research might, for instance, investigate the compatibility of the findings of this project with a network approach to gender assignment – in which gender is not assigned as a result of the application of rules/constraints, but nouns instead receive gender on the basis of patterns of associations between stored items (e.g. MacWhinney et al. 1989) – and determine how successfully such an approach is able to account for the German data in comparison to OGAT II.

A final area for further consideration is the application of the findings of the present study to the field of second language acquisition. As stated in 1.5, gender remains a significant problem for German learners and is one of the greatest sources of errors for even advanced learners of German (e.g. Heringer 1995, Wegera 1997, Lemhöfer et al. 2008, Bobb et al. 2015). It has, however, been demonstrated that explicit teaching of gender assignment regularities can lead to a reduction in gender errors among L2 learners (e.g. Schirrmeyer 2013, Kraiss 2014).

In his study, Kraiss (2014) in fact demonstrates the effectiveness of teaching Steinmetz's (1985, 1986, 2001, 2006) principles of "gender tally" and "gender eclipsis" to L2 learners as a method for learning German gender. As outlined in 1.4.5, these are the principles upon which Rice's (2006) theory of OGAT is based – gender tally referring to counting the number of times each gender is assigned to a noun by all applicable rules, and gender eclipsis referring to the application of a default gender hierarchy.

When comparing the results achieved in a series of tests taken by a test group, who were explicitly taught the gender tally/gender eclipsis mechanism, with those achieved by a control group, who received no explicit instruction on gender, it was found by Kraiss (2014) that the test group performed significantly better on the gender tests than the control group.

The application of OGAT II to the field of L2 acquisition thus seems promising. Initial areas of investigation might include the assessment of all GENDER FEATURES constraints in order to determine the order in which the constraints should be learnt. It has been claimed that this order should be dependent on two main factors: the number of exceptions a constraint has, such that the constraints with the highest percentage coverage of candidate nouns are learnt first (e.g. Mills 1986: 85, 109-116; Altmann & Raettig 1973: 302), and the scope or “reach” of a constraint, such that constraints with the highest number of candidate nouns (i.e. those which involve highly frequent features) are learnt first (e.g. Mills 1986: 85, 109-116; Heringer 1995: 212). The data from 2.3 can be used as a starting point for determining these.

In sum, the findings from this study suggest that OGAT II is a largely accurate representation of the way in which native German speakers assign gender to nouns, thus supporting the view that an area of the German language which has traditionally been viewed as notoriously opaque is in fact largely regular and systematic. It is hoped that future work on the topic will further demystify the process of grammatical gender assignment in German, not least for learners of German as a foreign language.

References

- Afonso, Olivia, Alberto Domínguez, Carlos J. Álvarez & David Morales. 2014. Sublexical and lexico-syntactic factors in gender access in Spanish. *Journal of Psycholinguistic Research* 43(1). 13-25.
- Aksenov, A.T. 1984. K probleme ekstralingvističeskoj motivacii grammatičeskoj kategorii roda [On the problem of the extralinguistic motivation for the grammatical category of gender]. *Voprosy jazykoznanija* 33. 14-25.
- Alcázar, Asier & Mario Saltarelli. 2014. *The syntax of imperatives*. Cambridge: Cambridge University Press.
- Allen, Margaret. 1978. *Morphological investigations*. Storrs, CT: University of Connecticut dissertation.
- Allerton, David John. 1979. *Essentials of grammatical theory: A consensus view of syntax and morphology*. London: Routledge & Kegan Paul.
- Altmann, Gabriel & Volker Raettig. 1973. Genus und Wortauslaut im Deutschen. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 26. 297-303.
- Apresjan, Ju D. 1974. Regular polysemy. *Linguistics* 14(2). 5-32.
- Arndt, Walter W. 1970. Nonrandom assignment of loanwords: German noun gender. *Word* 26. 244-253.
- Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. Cambridge, MA: MIT Press.
- Asher, Nicholas. 2011. *Lexical meaning in context: A web of words*. Cambridge: Cambridge University Press.
- Ashley, Karen C., Lynette Disch, David C. Ford, Erin MacSaveny, Steve Parker, Carla Unseth, Angela M. Williams, Rachel Wong & Brendon Yoder. 2010. How many constraints are there? A preliminary inventory of OT phonological constraints. *Graduate Institute of Applied Linguistics Dallas Occasional Papers in Applied Linguistics* 9. 1-36.
- Askedal, John Ole. 2016. Norwegian. In Peter Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds), *Word-formation: An international handbook of the languages of Europe*, Volume 4, 2525-2553. Berlin: de Gruyter.
- Audring, Jenny. 2016. Gender. *Oxford research encyclopaedias: Linguistics*. Oxford: Oxford University Press. <http://linguistics.oxfordre.com/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-43#acrefore-9780199384655-e-43-div1-1>. (15 May, 2018).
- Augst, Gerhard. 1975. *Untersuchungen zum Morpheminventar der deutschen Gegenwartssprache*. Tübingen: Narr.
- Augst, Gerhard. 1979. Neuere Forschungen zur Substantivflexion. *Zeitschrift für germanistische Linguistik* 7. 220-232.

- Avila, César, Matthew Lambon Ralph, Maria-Antònia Parcet, Daniel Geffner & José-Manuel Gonzalez-Darder. 2001. Implicit word cues facilitate impaired naming performance: Evidence from a case of anomia. *Brain and Language* 79. 185-200.
- Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baayen, R. Harald, Richard Piepenbrock & Leon Gulikers. 1995. The CELEX lexical database. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Bach, Emmon & Robert T. Harms. 1972. How do languages get crazy rules? In Robert Stockwell & Ronald Macaulay (eds), *Linguistic change and generative theory: Essays from the UCLA conference on historical linguistics in the perspective of transformational theory*, 1-21. Bloomington, IA: Indiana University Press.
- Badecker, William, Michele Miozzo & Raffaella Zanutti. 1995. The two-stage model of lexical retrieval: Evidence from a case of anomia with selective preservation of grammatical gender. *Cognition* 57. 193-216.
- Balteiro, Isabel. 2007. *The directionality of conversion in English*. Frankfurt am Main: Peter Lang.
- Bates, Elizabeth, Antonella Devescovi, Luigi Pizzamiglio, Simona D'Amico & Arturo Hernandez. 1995. Gender and lexical access in Italian. *Perception and Psychophysics* 57(6). 847-862.
- Batliner, Anton. 1984. The comprehension of grammatical and natural gender: a cross-linguistic experiment. *Linguistics* 22. 831-56.
- Bauch, Hans-Jürgen. 1971. Zum Informationsgehalt der Kategorie Genus im Deutschen, Englischen und Polnischen. *Wissenschaftliche Zeitschrift der Universität Rostock, Gesellschafts- und Sprachwissenschaftliche Reihe* 20(6). 411-419.
- Bauer, Laurie. 1983. *English word-formation*. Cambridge: Cambridge University Press.
- Bauer, Laurie. 1990. Be-heading the word. *Journal of Linguistics* 26. 1-31.
- Bauer, Laurie. 1992. Scalar productivity and -lily adverbs. In Geert Booij & Jaap van Marle (eds), *Yearbook of Morphology 1991*, 185-91. Dordrecht: Kluwer.
- Bauer, Laurie. 2008. Exocentric compounds. *Morphology* 18. 51-74.
- Bauer, Laurie. 2017. *Compounds and compounding*. Cambridge: Cambridge University Press.
- Beard, Robert. 2001. Derivation. In Andrew Spencer & Arnold M. Zwicky (eds), *The handbook of morphology*, 44-65. Oxford: Blackwell.
- Beito, Olav T. 1954. *Genusskifte i nynorsk*. Oslo: Jacob Dybwad.
- Beretta, Alan, Robert Fiorentino & David Poeppel. 2005. The effects of homonymy and polysemy on lexical access: An MEG study. *Cognitive Brain Research* 24(1). 57-65.
- Berkum, Jos J. A. van. 1996. *The psycholinguistics of grammatical gender*. Nijmegen: Nijmegen University Press.

- Berlin, Brent. 1972. Speculations on the growth of ethnobotanical nomenclature. *Language in Society* 1. 51-86.
- Berlin, Brent. 1978. Ethnobiological classification. In Eleanor Rosch & Barbara B. Lloyd (eds), *Cognition and categorization*, 9-26. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Berlin, Brent, Dennis Breedlove & Peter Raven. 1973. General principles of classification and nomenclature in folk biology. *American Anthropologist* 75. 212-242.
- Berlin, Brent, Dennis Breedlove & Peter Raven. 1974. *Principles of Tzeltal plant classification*. Cambridge, MA: Academic Press.
- Bierwisch, Manfred. 1983. Semantische und konzeptuelle Repräsentation lexikalischer Einheiten. In Rudolf Ružička & Wolfgang Motsch (eds), *Untersuchungen zur Semantik*, 61-99. Berlin: Akademie-Verlag.
- Bittner, Dagmar. 1994. Die Bedeutung der Genusklassifikation für die Organisation der deutschen substantivflexion. In Klaus-Michael Köpcke (ed.), *Funktionale Untersuchungen zur deutschen Nominal- und Verbmorphologie*, 65-80. Tübingen: Narr.
- Bittner, Dagmar. 2000. Gender classification and the inflectional system of German nouns. In Barbara Unterbeck & Matti Rissanen (eds), *Gender in grammar and cognition*, 1-24. Berlin: de Gruyter.
- Blank, Andreas. 2003. Polysemy in the lexicon and in discourse. In Brigitte Nerlich, Zazie Todd, Vimala Herman, David D. Clarke (eds), *Polysemy: Flexible patterns of meaning in mind and language*, 267-299. Berlin: de Gruyter.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt.
- Bobb, Susan, Judith F. Kroll & Carrie N. Jackson. 2015. Lexical constraints in second language learning: Evidence on grammatical gender in German. *Bilingualism* 18(3). 502-523.
- Boiger, Wolfgang. 2013. *Rechtskunde für RA-Fachangestellte*. Heidelberg: Müller.
- Booij, Geert. 1977. *Dutch morphology: A study of word formation in generative grammar*. Lisse: The Peter de Ridder Press/Dordrecht: Foris.
- Booij, Geert. 1992. Morphology, semantics and argument structure. In Iggy Roca (ed.), *Thematic Structure: Its relevance in grammar*, 48-64. Berlin: Foris.
- Booij, Geert. 2000. Inheritance. In Geert E. Booij, Christian Lehmann & Joachim Mugdan (eds), *Morphology. An international handbook of inflection and word formation*, Volume 1, 857-865. Berlin: de Gruyter.
- Booij, Geert. 2002. *The morphology of Dutch*. Oxford: Oxford University Press.
- Booij, Geert. 2005. *The grammar of words. An introduction to linguistic morphology*. Oxford: Oxford University Press.

- Bopp, Stephan. 2000-2018. Canoonet – Deutsche Grammatik, Online-Wörterbuch zu Rechtschreibung, Flexion und Wortbildung für die Sprache Deutsch. Universität Basel, Vrije Universiteit Amsterdam, IDSIA Lugano & Canoo Engineering AG. <http://www.canoo.net> (26 June, 2018).
- Braun, Friederike & Geoffrey Haig. 2010. When are German 'girls' feminine? How the semantics of age influences the grammar of gender agreement. In Markus Bieswanger, Heiko Motschenbacher & Susanne Muhleisen (eds), *Language in its socio-cultural context: New explorations in gendered, global and media uses*, 69-83. Frankfurt am Main: Peter Lang.
- Brown, Robert. 1958. How shall a thing be called? *Psychological Review* 65. 14-21.
- Brown, Robert. 1965. *Social psychology*. New York: Free Press.
- Brown, Susan Windisch. 2008. Polysemy in the mental lexicon. *Colorado Research in Linguistics* 21. 1-12.
- Brugman, Claudia. 1988. The story of over: Polysemy, semantics, and the structure of the lexicon. New York: Garland.
- Burkhard, Tanja. 2013. *Exploring Optimal Gender Assignment Theory for English loanwords in German*. Illinois: Southern Illinois University at Carbondale MA thesis.
- Butt, John & Carmen Benjamin. 2004. *A new reference grammar of modern Spanish*, 4th edn. London: Hodder Education.
- Bybee, Joan L. 1988. Morphology as lexical organization. In Michael Hammond & Michael Noonan (eds), *Theoretical morphology*, 119-141. Cambridge, MA: Academic Press.
- Caffarra, Sendy, Anna Siyanova-Chanturia, Francesca Pesciarelli, Francesco Vespignani & Cristina Cacciari. 2015. Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian. *Psychophysiology* 52(8) 1019-1030.
- Calabrese, Andrea. 2008. *Markedness and economy in a derivational model of phonology*. Berlin: de Gruyter.
- Callies, Markus, Eva Ogiermann & Konrad Szczesniak. 2010. Genusschwankung bei der Integration von englischen Lehnwörtern im Deutschen und Polnischen. In Carmen Scherer & Anke Holler (eds), *Strategien der Integration und Isolation nicht-nativer Einheiten und Strukturen*, 65-86. Berlin: de Gruyter.
- Callies, Markus, Alexander Onysko & Eva Ogiermann. 2012. Investigating gender variation of English loanwords in German. In Cristiano Furiassi, Virginia Pulcini & Félix Rodríguez González (eds), *The anglicization of European lexis*, 65-89. Amsterdam: John Benjamins.
- Caramazza, Alfonso & Ellen Grober. 1976. Polysemy and the structure of the subjective lexicon. In Clea Rameh (ed.), *Semantics: Theory and application*, 181-206. Washington: Georgetown University Round Table on Languages and Linguistics.
- Caramazza, Alfonso & Michele Miozzo. 1997. The relation between syntactic and phonological knowledge in lexical access: Evidence from the "tip-of-the-tongue" phenomenon. *Cognition* 64. 309-343.

- Carstensen, Broder. 1980. The gender of English loan-words in German. *Studia Anglica Posnaniensia* 12. 3-25.
- Carston, Robyn. 2012. Word meaning and concept expressed. *Linguistic Review* 29(4). 607-623.
- Casson, Ronald W. 1981. Folk classification: relativity and universality. In Ronald W. Casson (ed.), *Language, culture and cognition*, 75-91. New York: Macmillan.
- Chan, Sze-Mun. 2005. *Genusintegration. Eine systematische Untersuchung zur Genuszuweisung englischer Entlehnungen in der deutschen Sprache*. Munich: Iudicum.
- Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.
- Christen, Helen. 1998. Die Mutti oder das Mutti, die Rita oder das Rita? Über Besonderheiten der Genuszuweisung bei Personen- und Verwandtschaftsnamen in schweizerdeutschen Dialekten. In André Schnyder et al. (eds), *Ist mir getroumet mîn leben? Vom Träumen und vom Anderssein*. Festschrift für Karl-Ernst Geith zum 65. Geburtstag, 267-281. Göppingen.
- Clark, Eve E. 1993. *The lexicon in acquisition*. Cambridge: Cambridge University Press.
- Clyne, Michael. 1967. *Transference and triggering*. The Hague: Nijhoff.
- Comrie, Bernard. 1999. Grammatical gender systems: A linguist's assessment. *Journal of Psycholinguistic Research* 28(5). 457-466.
- Conzett, Philipp. 2006. Gender assignment and the structure of the lexicon. *Sprachtypol. Univ. Forsch. (STUF)* 59(3). 223-240.
- Copestake, Ann & Ted Briscoe. 1995. Semi-productive polysemy and sense extension. In James Pustejovsky & Branimir Boguraev (eds), *Lexical semantics: The problem of polysemy*, 15-67. New York: Clarendon Press.
- Corbett, Greville. 1979. The agreement hierarchy. *Journal of Linguistics* 15. 203-224.
- Corbett, Greville. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville. 1994. Gender and gender systems. In Ronald E. Asher, & James M. Y. Simpson (eds), *The encyclopedia of language and linguistics*, Volume 3, 1347-1353. Oxford: Pergamon Press.
- Corbett, Greville. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2013. Number of genders. In Matthew S. Dryer & Martin Haspelmath (eds), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/30>. (20 June, 2018).
- Corbett, Greville, Norman Fraser & Scott McGlashan (eds), 1993. *Heads in grammatical theory*. Cambridge: Cambridge University Press.
- Corbett, Greville & Norman Fraser. 2000. Default genders. In Barabara Unterbeck & Matti Rissanen (eds), *Gender in grammar and cognition*, 55-97. Berlin: de Gruyter.
- Croft, William & Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.

- Crowhurst, Megan. 2001. Coda conditions and um infixation in Toba Batak. *Lingua* 111. 561-590.
- Crowhurst, Megan & Lev Michael. 2005. Iterative footing and prominence driven stress in Nanti (Kampa). *Language* 81. 47-95.
- Cruse, Alan. 1986. *Lexical semantics*. Cambridge: Cambridge University Press.
- Cruse, Alan. 2000. *Meaning in language: An introduction to semantics and pragmatics*. Oxford: Oxford University Press.
- Dahl, Östen. 2000. Animacy and the notion of semantic gender. In Barbara Unterbeck & Matti Rissanen (eds), *Gender in grammar and cognition*, 99-115. Berlin: de Gruyter.
- Davenport, Mike & Hannahs, S. J. 2010. *Phonetics and phonology*. London: Routledge.
- Davies, Winifred V. 2012. Myths we live and speak by. Ways of imagining and managing language and languages. In Matthias Hüning, Ulrike Vogl & Olivier Moliner (eds), *Standard languages and multilingualism in European history*, 45-70. Amsterdam: John Benjamins.
- De Martino, Maria, Giulia Bracco & Alessandro Laudana. 2010. The activation of grammatical gender information in processing Italian nouns. *Language and Cognitive Processes* 26. 745-776.
- Deane, Paul D. 1988. Polysemy and cognition. *Lingua* 75(4). 325-361.
- Desrochers, Allan & Alain Paivio. 1990. Le phonem initial des noms inanimés et son effet sur l'identification du genre grammatical. *Canadian Journal of Psychology* 44. 44-57.
- Deutsches Anwalt Office Premium: Die digitale Fachbibliothek für die anwaltliche Praxis* – Haufe.de, 2013. https://www.haufe.de/recht/deutsches-anwalt-office-premium/1-einfuehrung-ii-das-voraus-nach-1932bgb_idesk_PI17574_HI8803493.html (8 September, 2017).
- Di Meola, Claudio & Daniela Puato. 2017. Wie viel grammatische Variation vertragen DaF-Lernende? In Claudio Di Meola, Joachim Gerdes, Livia Tonelli (eds), *Grammatik im fremdsprachlichen Deutschunterricht: Linguistische und didaktische Überlegungen zu Übungsgrammatiken*, 29-60. Berlin: Frank & Timme.
- Di Sciullo, Anna Maria & Edwin Williams. 1987. *On the definition of word*. Cambridge, MA: MIT Press.
- Dobrić, Nikola. 2014. On some problems of meaning – polysemy between sense enumeration and core meaning paradigms. *Filozofija I Društvo* 25(4). 146-163.
- Dölling, Johannes. 2018. Systematic polysemy. https://www.researchgate.net/publication/326083322_Systematic_Polysemy (24 June, 2018). To appear in Lisa Matthewson, Cécile Meier, Hotze Rullmann & Thomas Ede Zimmermann (eds), *The Blackwell Companion to Semantics*. Oxford: Wiley-Blackwell.
- Donaldson, Bruce. 2008. *Dutch: A comprehensive grammar*, 2nd edn. London: Routledge.
- Donalies, Elke. 2002. *Die Wortbildung des Deutschen*. Tübingen: Narr.
- Dressler Wolfgang U. 2006. Compound types. In Gonia Libben & Gary Jarema (eds), *The representation and processing of compound words*, 23-44. Oxford: Oxford University Press.

- Dressler, Wolfgang U. & Lavinia Merlini Barbaresi. 1994. *Morphopragmatics: Diminutives and intensifiers in Italian, German and other languages*. Berlin: de Gruyter.
- Dressler, Wolfgang U. & Anna Thornton. 1996. Italian nominal inflection. *Wiener Linguistische Gazette* 55-56. 1-24.
- Duden-Bibliothek Express*, version 6.0.2.3. 1993-2014. Berlin: Bibliographisches Institut.
- Duden – die deutsche Rechtschreibung*, 26th edn. 2013. Berlin: Bibliographisches Institut.
- Duden – Grammatik der deutschen Gegenwartssprache*, 8th edn. 2009. Berlin: Bibliographisches Institut.
- Duden Online-Wörterbuch*. 2018. Berlin: Bibliographisches Institut. www.duden.de/rechtschreibung. (20 June, 2018).
- Duden Online-Wörterbuch. 2018a. Aviso. www.duden.de/rechtschreibung/Aviso (10 July, 2018).
- Duden Online-Wörterbuch. 2018b. Band. www.duden.de/rechtschreibung/Band (10 July, 2018).
- Duden Online-Wörterbuch. 2018c. Firma. www.duden.de/rechtschreibung/Firma (10 July, 2018).
- Duden Online-Wörterbuch. 2018d. Korpuskel. www.duden.de/rechtschreibung/Korpuskel (17 July, 2018).
- Duden Online-Wörterbuch. 2018e. Leiter. www.duden.de/rechtschreibung/Leiter (10 July, 2018).
- Duden Online-Wörterbuch. 2018f. Virus. www.duden.de/rechtschreibung/Virus (10 July, 2018).
- Duden Sprachwissen. 2017. Die Verteilung der Artikel (Genusangabe) im Rechtschreibduden. <http://www.duden.de/sprachwissen/sprachratgeber/die-verteilung-der-artikel-genusangabe-im-rechtschreibduden>. (12 November, 2017).
- Duden – das Synonymwörterbuch*, 5th edn. 2010. Berlin: Bibliographisches Institut.
- Durrell, Martin. 2002. *Hammer's German grammar and usage*, 4th edn. London: Hodder Education.
- Durrell, Martin. 2011. *Hammer's German grammar and usage*, 5th edn. London: Hodder Education.
- DWB – Deutsches Wörterbuch*. 1998-2018. Trier Center for Digital Humanities. Universität Trier. woerterbuchnetz.de (13 July, 2018).
- DWDS – Das Digitale Wörterbuch der Deutschen Sprache des 20. Jahrhunderts*. Grötschel, Martin, Alexander Geyken & Lothar Lemnitzer. Berlin-Brandenburgische Akademie der Wissenschaften. www.dwds.de (8 August, 2018).
- Ebert, Robert Peter, Oskar Reichmann, Hans-Joachim Solms & Klaus-Peter Wegera. 1993. *Frühneuhochdeutsche Grammatik*. Tübingen: Max Niemeyer Verlag.
- Eisenberg, Peter. 1999. *Grundriss der deutschen Grammatik: Der Satz*, 3rd edn. Stuttgart: J.B. Metzler.

- Eisenberg, Peter. 2013. *Grundriss der deutschen Grammatik: Der Satz*, 4th edn. Stuttgart: J.B. Metzler.
- Elsaß, Robert & Stephan Möller. 2003a. Artikel+Vorname. *Atlas zur deutschen Alltagssprache*. www.atlas-alltagssprache.de/artikelvorname. (17 January, 2018).
- Elsaß, Robert & Stephan Möller. 2003b. Artikel+Nachname. *Atlas zur deutschen Alltagssprache*. www.atlas-alltagssprache.de/artikelnachname. (17 January, 2018).
- Elsaß, Robert & Stephan Möller. 2003c. Artikelform. *Atlas zur deutschen Alltagssprache*. www.atlas-alltagssprache.de/artikelform. (17 January, 2018).
- Elst, Gaston van der. 1983. Abweichung bei der Genusordnung in den deutschen Dialekten. In Werner Besch et al. (eds), *Dialektologie: Ein Handbuch zur deutschen und allgemeinen Dialektforschung. Halbband 2*, 1202-1209. Berlin: de Gruyter.
- El-Youssef, Nadia. 2010. Gender assignment to new words in Catalan. Barcelona: Universitat Autònoma de Barcelona MA thesis.
- Engel, Ulrich. 2009. *Deutsche Grammatik*, 2nd edn. Munich: Iudicium.
- Enger, Hans-Olav. 2007. Grammaticalisation due to homonymy avoidance in Gudbrandsdalen /æpr̥ede/? *Transactions of the Philological Society* 105. 42-65.
- Enger, Hans-Olav. 2009. The role of core and non-core semantic rules in gender assignment. *Lingua* 119. 1281-1299.
- Evans, Vyvyan. 2004. How we conceptualise time: Language, meaning and temporal cognition. *Essays in Arts and Sciences* 33(2). 13-44.
- Evans, Vyvyan. 2005. The meaning of time: polysemy, the lexicon and conceptual structure. *Journal of Linguistics* 41(1). 33-75.
- Evans, Vyvyan & Melanie Green. 2006. *Cognitive linguistics: An introduction*. Edinburgh: Edinburgh University Press.
- Falkum, Ingrid. 2011. *The semantics and pragmatics of polysemy: A relevance-theoretic account*. London: University College London dissertation.
- Falkum, Ingrid & Agustín Vicente. 2015. Polysemy: Current perspectives and approaches. *Lingua* 157. 1-16.
- Fehring, Carol. 1999. *A reference grammar of Dutch*. Cambridge: Cambridge University Press.
- Fischer, Rudolf-Josef. 2005. *Genuszuordnung. Theorie und Praxis am Beispiel des Deutschen*. Frankfurt am Main: Peter Lang.
- Flämig, Walter. 1991. *Grammatik des Deutschen: Einführung in Struktur- und Wirkungszusammenhänge*. Berlin: Akademie-Verlag.
- Fleischer, Wolfgang & Irmhild Barz. 1992. *Wortbildung der deutschen Gegenwartssprache*. Tübingen: Niemeyer.

- Fleischer, Wolfgang & Irmhild Barz. 2012. *Wortbildung der deutschen Gegenwartssprache*, 4th edn. Berlin: de Gruyter.
- Fox, Anthony. 2005. *The structure of German*. Oxford: Oxford University Press.
- Franco, Karlien, Eline Zenner & Dirk Speelman. 2018. Let's agree to disagree. (Variation in) the assignment of gender to nominal Anglicisms in Dutch. *Journal of Germanic Linguistics* 30. 43-87.
- Frazier, Lyn & Keith Rayner. 1990. Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language* 29(2). 181-200.
- Frieke, F. C. 1988. De voorspelbaarheid van 'de' en 'het' eenlettergrepige nomina in het Nederlands. Amsterdam: VU Amsterdam dissertation.
- Fries, Norbert. 1997. *Die hierarchische Organisation grammatischer Kategorien*. Debrecen: Kossuth Egyetemi Kiadó.
- Frisson, Stephen. 2009. Semantic underspecification in language processing. *Language and Linguistics Compass* 3(1). 111-127.
- Fuller, Janet & Lehnert, Heike. 2000. Noun phrase structure in German-English codeswitching: variation in gender assignment and article use. *International Journal of Bilingualism* 4(3). 399-420.
- Gagné, Christina & Thomas Spalding. 2006. Conceptual combinations: Implications for the mental lexicon. In: Gonia Libben & Gary Jarema (eds), *The representation and processing of compound words*, 145-168. Oxford: University Press.
- Galbreath, Blake L. E. 2010. Gender assignment in Contemporary Standard Russian: A comprehensive analysis in Optimality Theory. Washington: Washington State University dissertation.
- Gärdenfors, Peter. 1999. Some tenets of cognitive semantics. In Jens Allwood & Peter Gärdenfors (eds), *Cognitive semantics: Meaning and cognition*, 19-36. Amsterdam: John Benjamins.
- Geeraerts, Dirk. 1993. Vagueness's puzzles, polysemy's vagaries. *Cognitive Linguistics* 4(3). 223-72.
- Geerts, Guido. 1975. Het genus van Engelse leenwoorden in het Duits en in het Nederlands. In Ria Jansen-Sieben, S. de Vriendt & Roland Willemyns (eds), *Spel van zinnen: Album A. van Loey*, 115-123. Brussels: Ed. de l'Univ.
- Geerts, Guido. 1996. De genusbepalende eigenschappen van Engelse leenwoorden in het Nederlands. *Mededelingen van de Koninklijke Academie voor Nederlandse Taal- en Letterkunde* 106. 137-146.
- Gollan, Tamar H. & Ram Frost. 2001. Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research* 30. 627 - 651.
- Graedler, Anne-Line. 2013. A corpus-based study of gender assignment in recent English loanwords in Norwegian. In Hilde Hasselgård, Jarle Ebeling & Signe Oksefjell Ebeling (eds), *Corpus perspectives on patterns of lexis*, 209-222. Amsterdam: John Benjamins.

- Gregor, Bernd. 1983. *Genuszuordnung: Das Genus englischer Lehnwörter im Deutschen*. Tübingen: Niemeyer.
- Greisbach, Reinhold. 2003. The pronunciation of loanwords in German. *The 15th International Congress of Phonetic Sciences (ICPhS)*. 799-802.
- Guevara, Emiliano & Sergio Scalise. 2009. Searching for universals in compounding. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds), *Universals of language today*, 101-128. Berlin: Springer.
- Guy, Gregory. 2011. Variability. In Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume & Keren Rice (eds), *The Blackwell companion to phonology*, Volume 4, 2190-2213. Oxford: Wiley-Blackwell.
- Haas, Wim de & Mieke Trommelen. 1993. *Morfologisch handboek van het Nederlands. Een overzicht van de woordvorming*. 's-Gravenhage: SDU Uitgeverij.
- Haeseryn, Walter, Kirsten Romijn, Guido Geerts, Jaap de Rooij & Maarten van de Toorn. 1997. *Algemene Nederlandse Spraakkunst*, 2nd edn. Groningen/Deurne: Martinus Nijhoff/Wolters.
- Hamburger Abendblatt*. 07.08.08. New York, USA. <https://www.abendblatt.de/hamburg/article/107436172/New-York-USA.html> (20 January, 2018).
- Hamm, Fritz & Hans Kamp. 2009. Ontology and inference: The case of German *ung*-nominals. In Antje Roßdeutscher (ed.), *Disambiguation and reambiguation*, 1-67. Stuttgart: University of Stuttgart.
- Henaff Gonon, Marie-Anne, Remy Bruckert & Francois Michel. 1989. Lexicalization in an anomic patient. *Neuropsychologia* 27. 391-407.
- Hernandez, Arturo E., Sonja A Kotz, Juliane Hofmann, Vivian V. Valentin, Mirella Dapretto & Susan Y. Bookheimer. 2004. The neural correlates of grammatical gender decision in Spanish. *Neuroreport* 15. 863-866.
- Harnisch, Rüdiger & Damaris Nübling. 2004. Namenkunde. In Geert E. Booij, Christian Lehmann, Joachim Mugdan & Stavros Skopeteas (eds), *Morphologie*, Volume 2, 1901-1910. Berlin: de Gruyter.
- Haspelmath, Martin & Andrea Sims. 2010. *Understanding morphology*, 2nd edn. London: Hodder.
- Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and case studies*. Chicago: University of Chicago Press.
- Hayes, Bruce. 2004. Phonetically driven phonology: The role of Optimality Theory and Inductive Grounding. In John J. McCarthy (ed.), *Optimality Theory in phonology: A reader*, 290-309. Oxford: Blackwell.
- Heringer, Hans Jürgen. 1995. Prinzipien der Genuszuweisung. In Heidrun Popp (ed.), *Deutsch als Fremdsprache: An den Quellen eines Faches. Festschrift für Gerhard Helbig zum 65. Geburtstag*, 203-216. Munich: Iudicium.
- Hickey, Raymond. 2000. On the phonology of gender in Modern German. In Matti Rissanen & Barbara Unterbeck (eds), *Gender in grammar and cognition*, 621-633. Berlin: de Gruyter.

- Hjelde, Arnstein. 1996. The gender of English nouns used in American Norwegian. In P. Sture Ureland & Iain Clarkson (eds), *Language contact across the North Atlantic*, 297-312. Tübingen: Niemeyer.
- Hoberg, Ursula. 2004. *Grammatik des Deutschen im europäischen Vergleich. Das Genus des Substantivs*. Mannheim: Institut für Deutsche Sprache. (amades. Arbeitspapiere und Materialien zur deutschen Sprache. 3/04).
- Hock, Hans Heinrich & Brian Joseph. 2009. *Language history, language change, and language relationship. An introduction to historical and comparative linguistics*. Berlin: de Gruyter.
- Hockett, Charles. F. 1958. *A course in modern linguistics*. New York: Macmillan.
- Hoeksema, Jakob. 1986. *Categorial morphology*. New York: Garland Press.
- Hoeksema, Jacob. 1992. The head parameter in morphology and syntax. In Gilbers, Dicky & Looyenga, Sietze (eds), *Language and cognition: Yearbook of the research group for linguistic theory and knowledge of the University of Groningen*, Volume 2. Groningen: Universiteitsdrukkerij Groningen. 119-132.
- Hohlfeld, Annette. 2006. Accessing grammatical gender in German: The impact of gender-marking regularities. *Applied Psycholinguistics* 27. 127-142.
- Hornstein, Norbert, Jairo Nunes & Kleantes K. Grohmann. 2005. *Understanding minimalism*. Cambridge: Cambridge University Press.
- Hudson, Richard. A. 1987. Zwicky on heads. *Journal of Linguistics* 23. 109-132.
- Hüning, Matthias. 2008. Adjective + Noun constructions between syntax and word formation in Dutch and German. In Sascha Michel & Alexander Onysko (eds), *Cognitive perspectives on word formation*, 195-215. Berlin: de Gruyter.
- Hunt, Jaime. 2016. The gender of Anglicisms in German: Towards a predictive model. Talk presented at the Forum for Germanic Language Studies, 7-8 January 2016, Aston, UK. https://www.researchgate.net/publication/295186020_The_gender_of_anglicisms_in_German_Towards_a_predictive_model (20 January, 2018).
- Iacobini, Claudio. 2000. Base and direction of derivation. In Geert E. Booij, Christian Lehmann & Joachim Mugdan (eds), *Morphology. An international handbook of inflection and word formation*, Volume 1, 865-876. Berlin: de Gruyter.
- Jescheniak, Jörg D. & Willem J. M. Levelt. 1994. Word frequency effects in speech production. Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory & Cognition* 20(4). 824-843.
- Jesenšek, Vida. 1995. Medienwirksame Neologismen in der deutschen Gegenwartssprache. *Linguistica* 2. 171-207.
- Johnson, Keith. 2008. *Quantitative methods in linguistics*. Oxford: Blackwell.
- Kaeding, Friedrich Wilhelm. 1897/1898. *Häufigkeitwörterbuch der deutschen Sprache. Festgestellt durch einen Arbeitsausschuß der deutschen Stenographiesysteme*. Berlin: Kaeding.

- Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
- Kageyama, Taro. 2008. Semantic effects of left-hand elements on right-hand head structure. *136th Linguistic Society of Japan*. 68-73.
- Kageyama, Taro. 2009. Isolate: Japanese. In Rochelle Lieber & Pavol Štekauer (eds), *The Oxford handbook of compounding*, 512-526. Oxford: Oxford University Press.
- Karmiloff-Smith, Annette. 1979. *A functional approach to child language: A study of determiners and reference*. Cambridge: Cambridge University Press.
- Kastovsky, Dieter. 2005. Conversion and/or zero: word-formation theory, historical linguistics, and typology. In Laurie Bauer & Salvador Valera Hernández (eds), *Approaches to conversion/zero-derivation*, 31-50. Münster/New York: Waxmann.
- Kastovsky, Dieter. 2006. Morphology as word formation in 20th century linguistics. In Sylvain Auroux, E.F.K. Koerner, Hans-Josef Niederehe, Kees Versteegh (eds), *History of the language sciences: An international handbook on the evolution of the study of language from the beginnings to the present, 2324-2339*. Berlin: de Gruyter.
- Katz, Jerrold J. 1972. *Semantic Theory*. New York: Harper & Row.
- Katz, Jerrold J. & Jerry A. Fodor. 1963. The structure of a semantic theory. *Language* 39(2). 170-210.
- Katz, Jerrold J. & Paul M. Postal. 1964. *An integrated theory of linguistic descriptions*. Cambridge, MA: MIT Press.
- Kennedy, Benjamin Hall. 2010. *The revised Latin primer*. Cambridge: Cambridge University Press.
- Kiparsky, Paul. 1982. *Explanation in phonology*. Dordrecht: Foris.
- Klein, Devorah E. & Gregory L. Murphy. 2001. The representation of polysemous words. *Journal of Memory and Language* 45(2). 259-282.
- Klein, Devorah E. & Gregory L. Murphy. 2002. Paper has been my ruin: conceptual relations of polysemous senses. *Journal of Memory and Language* 47(4). 548-570.
- Klepousniotou, Ekaterini. 2002. The processing of lexical ambiguity: Homonymy and polysemy in the mental lexicon. *Brain and Language* 81(1-3). 205-223.
- Klepousniotou, Ekaterini. 2007. Reconciling linguistics and psycholinguistics: On the psychological reality of linguistic polysemy. In Marina Rakova, Gergely Petho & Csilla Rakosi (eds), *The cognitive basis of polysemy: New sources of evidence for theories of word meaning*, 17-46. Frankfurt am Main: Peter Lang.
- Klepousniotou, Ekaterini & Shari R. Baum. 2007. Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics* 20(1). 1-24.
- Klepousniotou, Ekaterini, Debra Titone & Carolina Romero. 2008. Making sense of word senses: the comprehension of polysemy depends on sense overlap. *Journal of Experimental Psychology: Learning, Memory and Cognition* 34(6). 1534-1543.

- Kluge, Friedrich. 2002. *Etymologisches Wörterbuch der deutschen Sprache*, 24th edn. Berlin: de Gruyter.
- König, Werner. 2001. *Dtv-Atlas Deutsche Sprache*. Munich: DTV.
- Köpcke, Klaus-Michael. 1982. *Untersuchungen zum Genusystem der deutschen Gegenwartssprache*. Tübingen: Niemeyer.
- Köpcke, Klaus-Michael & David Zubin. 1983. Die kognitive Organisation der Genuszuweisung zu den einsilbigen Nomen der deutschen Gegenwartssprache. *Zeitschrift für germanistische Linguistik* 11(2). 166-182.
- Köpcke, Klaus-Michael & David Zubin. 1984. Sechs Prinzipien für die Genuszuweisung im Deutschen: Beitrag zur natürlichen Klassifikation. *Linguistische Berichte* 93. 26-50.
- Köpcke, Klaus-Michael & David Zubin. 1996. Prinzipien für die Genuszuweisung im Deutschen. In Ewald Lang & Gisela Zifonun (eds), *Deutsch - typologisch. Jahrbuch des Instituts für Deutsche Sprache*, 473-491. Berlin: de Gruyter.
- Köpcke, Klaus-Michael & David Zubin. 1997. Sechs Prinzipien für die Genuszuweisung im Deutschen: Beitrag zur natürlichen Klassifikation. In Heinz Sieburg (ed.), *Sprache – Genus/Sexus*, 86-115. Frankfurt: Lang.
- Köpcke, Klaus-Michael & David Zubin. 2005a. Nominalphrasen ohne lexicalischen Kopf – Zur Bedeutung des Genus für die Organisation des mentalen Lexikons am Beispiel der Autobezeichnungen im Deutschen. *Zeitschrift für Sprachwissenschaft* 24. 93-122.
- Köpcke, Klaus-Michael & David Zubin. 2005b. Metonymic pathways to neuter-gender human nominals. In Klaus Panther & Linda Thornburg (eds), *Metonymy and pragmatic inferencing*, 149-166. Amsterdam: John Benjamins.
- Köpcke, Klaus-Michael & David Zubin. 2009. Genus. In Elke Hentschel & Petra M. Vogel (eds), *Deutsche Morphologie*, 132-154. Berlin: de Gruyter.
- Köpcke, Klaus-Michael, Klaus-Uwe Panther & David Zubin. 2010. Motivating grammatical and conceptual agreement in German. In Hans-Jörg Schmid & Susanne Handl (eds), *Cognitive foundations of linguistic usage patterns*, 171-194. Berlin: de Gruyter.
- Kraiss, Andrew. 2014. Taking the awful out of the German language: A study of a new way to teach German gender and plural assignment. *Die Unterrichtspraxis/Teaching German* 47(1). 23-41.
- Kratochvilová, Iva. 2000. Zur Problematik der Genuszuordnung englischer Entlehnungen in der deutschen Gegenwartssprache. *Brünner Beiträge zur Germanistik und Nordistik* 14. 67-75.
- Kripke, Saul A. 1980. *Naming and necessity*. Cambridge, MA: Harvard University Press.
- Kürschner, Sebastian & Damaris Nübling. 2011. The interaction of gender and declension in Germanic languages. *Folia Linguistica* 45(2). 355-388.
- Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: The University of Chicago Press.

- Lang, Adrienne. 1976. The semantic base of gender in German. *Lingua* 40. 55-68.
- Langacker, Ronald W. 1987. *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 1988. A usage-based model. In Brygida Rudzka-Ostyn (ed.), *Topics in cognitive linguistics*, 127-161. Amsterdam: John Benjamins.
- Langendonck, Willy van & Mark van der Velde. 2009. The functions of (in)definiteness: Markers with proper names. In Wolfgang Ahrens, Sheila Embleton, André et Lapierre (eds), *Names in multi-lingual and multi-ethnic contact. Proceedings of the 23rd ICOS congress*, 1005-1011. Toronto: York University.
- Lass, Roger. 1984. *Phonology: an introduction to basic concepts*. Cambridge: Cambridge University Press.
- Laurence, Stephen & Eric Margolis. 1999. Concepts and cognitive science. In Eric Margolis & Stephen Laurence (eds), *Concepts: Core readings*, 3-81. Cambridge, MA: MIT Press.
- Lefebvre, Claire. 2014. *Relabelling in language genesis*. Oxford: Oxford University Press.
- Leipzig Corpora Collection – Wortschatz Universität Leipzig*. 1998-2018. Institute of Computer Science at Leipzig University. <http://wortschatz.uni-leipzig.de/de>.
- Leiss, Elisabeth. 2005. Derivation als Grammatikalisierungsbrücke für den Aufbau von Genusdifferenzierungen im Deutschen. In Torsten Leuschner, Tanja Mortelmans & Sarah de Groot (eds), *Grammatikalisierung im Deutschen*, 11-30. Berlin: de Gruyter.
- Lemhöfer Kristin, Katharina Spalek & Herbert Schriefers. 2008. Cross-language effects of grammatical gender in bilingual word recognition and production. *Journal of Memory and Language* 59. 312-330.
- Lev-Ari, Shiri, Marcela San Giacomo & Sharon Peperkamp. 2014. The effect of domain prestige and interlocutors' bilingualism on sound adaptation. *Journal of Sociolinguistics* 18(5). 658-684.
- Levy, Erika S., Mira Goral & Loraine K. Obler. 2006. Doghouse/Chien-maison/Niche: Approaches to the understanding of compound processing in bilinguals. In Gonia Libben & Gary Jarema (eds), *The representation and processing of compound words*, 125-144. Oxford: Oxford University Press.
- Lewandowska-Tomaszczyk, Barbara. 2007. Polysemy, prototypes, and radial categories. In Dirk Geeraerts & Hubert Cuyckens (eds), *The Oxford handbook of cognitive linguistics*, 139-169. Oxford: Oxford University Press.
- Lieber, Rochelle. 1980. *On the organization of the lexicon*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Lieber, Rochelle. 1981. Morphological conversion within a restrictive theory of the lexicon. In Michael Moortgat, Harry van de Hulst & Teun Hoekstra (eds), *The scope of lexical rules. Linguistic models*, 161-200. Dordrecht: Foris.
- Lieber, Rochelle. 1989. On percolation. In Geert Booij & Jaap Van Marle (eds), *Yearbook of morphology 1989*, 95-138. Dordrecht: Kluwer.

- Lieber, Rochelle. 1992. *Deconstructing morphology: word formation in syntactic theory*. Chicago: University of Chicago Press.
- Löbner, Sebastian. 2002. *Semantik*. Berlin: de Gruyter.
- Lohde, Michael. 2006. *Wortbildung des modernen Deutschen: ein Lehr- und Übungsbuch*. Tübingen: Narr.
- Lohr, Clarissa. 2016. Do German speakers of non-binary genders use *es* and its declensions as their pronoun? www.quora.com/Do-German-speakers-of-non-binary-genders-use-es-and-its-declensions-as-their-pronoun (10 August, 2017).
- Lohr, Clarissa. 2017. Do non-binary German speakers ever introduce themselves using *Hallo ich bin das* (their name)? www.quora.com/Do-non-binary-German-speakers-ever-introduce-themselves-using-Hallo-ich-bin-das-___-their-name (10 August, 2017).
- Lüger, Heinz-Helmut. 1995. *Pressesprache*. Berlin: de Gruyter.
- MacKay, Donald G. 1999. Gender in English, German, and other languages: Problems with the old theory, opportunities for the new. In Ursula Pasero & Friederike Braun (eds), *Perceiving and performing gender: Herstellung und Wahrnehmung von Geschlecht*, 73-87. Wiesbaden: Westdeutscher Verlag.
- MacWhinney, Brian. 1978. The acquisition of morphophonology. *Monographs of the Society for Research in Child Development* 43. 1-123.
- MacWhinney, Brian, Jared Leinbach, Roman Taraban & Janet McDonald. 1989. Language learning: Cues or rules? *Journal of Memory and Language* 28. 255-277.
- Maratsos, Michael. 1979. Learning how and when to use pronouns and determiners. In Paul Fletcher & Michael Garman (eds), *Language acquisition*, 225-240. Cambridge: Cambridge University Press.
- Marchand, Hans. 1963. On a question of contrary analysis with derivationally connected but morphologically uncharacterized words. In Dieter Kastovsky (ed.), *Studies in syntax and word-formation, selected articles by Hans Marchand*, 224-241. International Library of General Linguistics. Munich: Fink.
- Marchand, Hans. 1964. A set of criteria for the establishing of derivational relationship between words unmarked by derivational morphemes. *Indogermanische Forschungen* 69. 10-19.
- Marchand, Hans. 1969. *The categories and types of present-day English word-formation. A synchronic-diachronic approach*. Munich: C. H. Beck'sche Verlagsbuchhandlung.
- Mastropavlou, Maria & Ianthi Tsimpli. 2011. The role of suffixes in grammatical gender assignment in Modern Greek: A psycholinguistic study. *Journal of Greek Linguistics* 11. 27-55.
- Matthews, Peter. 1974. *Morphology*. Cambridge: Cambridge University Press.
- Mayerthaler, Willi. 1981. *Morphologische Natürlichkeit*. Wiesbaden: Athenaion.
- McCarthy, John J. 2002. *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.

- McCarthy, John J. 2007. What is Optimality Theory? https://works.bepress.com/john_j_mccarthy/ (21 April, 2016).
- McCarthy, John J. 2008. *Doing Optimality Theory: Applying theory to data*. Oxford: Blackwell.
- McClelland, James L. & Jeffrey L. Elman. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18. 1-86.
- McMahon, April. 2000. *Change, chance, and optimality*. Oxford: Oxford University Press.
- Meier, Helmut. 1964. *Deutsche Sprachstatistik*. Hildesheim: Olms.
- Mel'čuk, Igor A. 1974. Statistics and the relationship between gender of French nouns and their endings. In V. Ju. Rozencvejk (ed.), *Essays on lexical semantics*, 11-42. Stockholm: Skriptor.
- Menzel, Barbara. 2004. *Genuszuweisung im DaF-Erwerb: Psycholinguistische Prozesse und didaktische Implikationen*. Berlin: Weißensee Verlag.
- Mervis, Carolyn & Eleanor Rosch. 1981. Categorization of natural objects. *Annual Review of Psychology* 32. 89-115.
- Mills, Anne. 1986. *The acquisition of gender: A study of German and English*. Berlin: Springer.
- Miozzo Michele & Alfonso Caramazza. 1997. Retrieval of lexical-syntactic features in tip-of-the-tongue states. *Journal of Experimental Psychology: Learning, Memory and Cognition* 23(6). 1410-1423.
- Mörth Karlheinz & Wolfgang U. Dressler. 2014. German plural doublets with and without meaning differentiation. In Franz Rainer, Francesco Gardani, Hans Christian Luschutzky, Wolfgang U. Dressler (eds), *Morphology and meaning: Selected papers from the 15th International Morphology Meeting, Vienna, February 2012*, 249-258. Amsterdam: John Benjamins.
- Muhvić-Dimanovski, Vesna. 1995. Anglicisms in German: the problem of variants. *Studia romanica et anglica Zagrabiana* 42. 269-274.
- Müller, Birgit. 2011. Kampagne gegen Leerstand und Wohnungslosigkeit. *Hinz&Kunzt* 217 (01.03.11) <https://www.hinzundkunzt.de/gegen-leerstand-wohnungslos/> (20 January, 2018).
- Müller, Gereon. 1999. Optionality in optimality-theoretic syntax. *GLoT International* 4(5). 3-8.
- Murphy, M. Lynne. 2010. *Lexical meaning*. Cambridge: Cambridge University Press.
- Nagy, Naomi & Bill Reynolds. 1997. Optimality Theory and variable word-final deletion in Faetar. *Language Variation and Change* 9. 37-55.
- Namiki, Takayasu. 2001. Further evidence in support of the right-hand head rule in Japanese. In Jeroen van de Weijer & Tetsuo Nishihara (eds), *Issues in Japanese phonology and morphology*, 277-297. Berlin: de Gruyter.
- Naumann, Bernd & Petra M. Vogel. 2000. Derivation. In Geert E. Booij, Christian Lehmann & Joachim Mugdan (eds), *Morphologie: Ein Internationales Handbuch zur Flexion und Wortbildung*, Volume 1, 929-942. Berlin: de Gruyter.

- Neef, Martin. 1998. A case study in declarative morphology: German case inflection. In Richard Wiese & Wolfgang Kehrein (eds), *Phonology and morphology of the Germanic Languages*, 219-240. Tübingen: Niemeyer.
- Neef, Martin. 2009. IE, Germanic: German. In Rochelle Lieber & Pavol Štekauer (eds), *The Oxford handbook of compounding*, 386-399. Oxford: Oxford University Press.
- Nelson, Donald. 1998. A prolegomena to a German gender dictionary. *Word* 49(2). 205-224.
- Nesset, Tore. 2003. Gender assignment in Ukrainian: Language specific rules and universal principles. *Poljarnyj Vestnik* 6. 71-85.
- Nesset, Tore. 2005. Rule counting vs. rule ordering: universal principles of rule interaction in gender assignment. In Geert Booij, Emiliano Guevara, Angela Ralli, Salvatore Sgroi & Sergio Scalise (eds), *Morphology and linguistic typology. On-line proceedings of the fourth mediterranean morphology meeting (MMM4) Catania, 21-23 September 2003*. University of Bologna.
- Nesset, Tore. 2006a. Gender meets the usage-based model: Four principles of rule interaction in gender assignment. *Lingua* 116. 1369-1393.
- Nesset, Tore. 2006b. Meta-constraints: Constraint interaction and gender assignment in Ukrainian. *Proceedings of the Thirtieth Annual Meeting of the Berkeley Linguistics Society*. 318-329.
- Ní Chiosáin, Máire. 1999. Syllables and phonotactics in Irish. In Harry van der Hulst & Nancy A. Ritter (eds), *The syllable: views and facts*, 551-575. Berlin: de Gruyter.
- Nordwest Zeitung Online*. 20.10.2017. Der Herr im Haus heißt Iran. www.nwzonline.de/politik/jerusalem-analyse-der-herr-im-haus-heisst-iran_a_32,1,562186008 (20 January, 2018).
- Norrick, Neal R. 1981. *Semiotic principles in semantic theory*. Amsterdam: John Benjamins.
- Nübling, Damaris. 2008. Was tun mit Flexionsklassen? Deklinationsklassen und ihr Wandel im Deutschen und seinen Dialekten. *Zeitschrift für Dialektologie und Linguistik* 75(3). 282-330.
- Nübling, Damaris. 2011. How do exceptions arise? On different paths to morphological irregularity. In Horst Simon & Heike Wiese (eds), *Expecting the unexpected: Exceptions in grammar* (Trends in Linguistics 216), 139-162. Berlin: de Gruyter.
- Nübling, Damaris. 2014. *Das Merkel* – Das Neutrum bei weiblichen Familiennamen als derogatives Genus? In Friedhelm Debus, Rita Heuser & Damaris Nübling (eds), *Linguistik der Familiennamen*, 205-232. Hildesheim: Olms.
- Nübling, Damaris. 2015. Die Bismarck – der Arena – das Adler. Vom Drei-Genus- zum Sechs-Klassen-System bei Eigennamen im Deutschen: Degrammatikalisierung und Exaptation. *Zeitschrift für germanistische Linguistik* 43(2). 306-344.
- Nübling, Damaris, Simone Busley & Juliane Drenda. 2013. *Dat Anna* und *s Eva* – Neutrale Frauenrufnamen in deutschen Dialekten und im Luxemburgischen zwischen pragmatischer und semantischer Genuszuweisung. *Zeitschrift für Dialektologie und Linguistik* 80(2). 152-196.
- Nunberg, Geoffrey. 1979. The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy* 3(2). 143-184.

- Nunberg, Geoffrey. 1995. Transfers of meaning. *Journal of Semantics* 12(2). 109-132.
- O'Donnell, Timothy J. 2015. *Productivity and reuse in language: A theory of linguistic computation and storage*. Cambridge, MA: MIT Press.
- Oehler, Heinz. 1966. *Grundwortschatz Deutsch*. Stuttgart: Klett.
- Olsen, Susan. 1986. *Wortbildung im Deutschen. Eine Einführung in die Theorie der Wortstruktur*. Stuttgart: Alfred Kröner Verlag.
- Olsen, Susan. 1991. Die deutsche Nominalphrase als 'Determinansphrase'. In Susan Olsen & Gisbert Fanselow (eds), *DET, COMP und INFL: Zur Syntax funktionaler Kategorien und grammatischer Funktionen*, 35-56. Tübingen: Niemeyer.
- Onysko, Alexander. 2007. *Anglicisms in German: Borrowing, lexical productivity and written codeswitching*. Berlin: de Gruyter.
- OWID – *Online-Wortschatz-Informationssystem Deutsch*. 2008-2010. Mannheim: Institut für Deutsche Sprache. www.owid.de (8 August, 2018).
- Oxford-Duden Pictorial German Dictionary*, 2nd edn. 1994. Oxford: Oxford University Press.
- Padovani, Roberto & Cristina Cacciari. 2003. Il ruolo della trasparenza morfologica nel riconoscimento di parole in Italiano [The role of morphological transparency in the recognition of Italian words]. *Giornale Italiano de Psicologia* 4. 749-771.
- Penke, Martina & Marion Krause. 2002. German noun plurals: A challenge to the dual-mechanism model. *Brain and Language* 81. 303-311.
- Pfau, Roland. 2009. *Grammar as processor: a distributed morphology account of spontaneous speech errors*. Amsterdam: John Benjamins.
- Pfeifer, Wolfgang. 2010. *Etymologisches Wörterbuch des Deutschen*. Hamburg: Akzente Verlagshaus.
- Pickering, Martin J. & Steven Frisson. 2001. Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27(2). 556-573.
- Pinker, Stephen & Alan Prince. 1991. Regular and irregular morphology and the psychological status of rules of grammar. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*. 230-251.
- Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivation*. Berlin: de Gruyter.
- Plag, Ingo. 2003. *Word-formation in English*. Cambridge: Cambridge University Press.
- Plank, Frans. 1986. Das Genus der deutschen Ge-Substantive und Verwandtes. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 39. 44-60.
- Popadić, Hanna. 1971. *Untersuchungen zur Frage der Nominalisierung des Verbalausdrucks im heutigen Zeitungsdeutsch*. Tübingen: Narr.

- Poplack, Shana, Alicia Pousada & David Sankoff. 1982. Competing influences on gender assignment: Variable process, stable outcome. *Lingua* 57(1). 1-28.
- Poplack, Shana & David Sankoff. 1984. Borrowing: the synchrony of integration. *Linguistics* 22. 99-135.
- Poplack, Shana, David Sankoff & Chris Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26(1). 47-104.
- Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Ms. Rutgers University and University of Colorado at Boulder.
- Prince, Alan & Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford/Malden/Carlton: Blackwell.
- Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, MA: MIT Press.
- Putnam, Hilary. 1975. The meaning of "meaning". *Minnesota Studies in the Philosophy of Science* 7. 131-193.
- Pylkkänen, Liina, Rodolfo Llinás & Gregory L. Murphy. 2006. The representation of polysemy: MEG evidence. *Journal of Cognitive Neuroscience* 18(1). 97-109.
- Quine, Willard van Orman. 1960. *Word and object*. Cambridge, MA: MIT Press.
- Radford, Andrew, Martin Atkinson, David Britain, Harald Clahsen & Andrew Spencer. 1999. *Linguistics: An introduction*. Cambridge: Cambridge University Press.
- Rainer, Franz. 1993. *Spanische Wortbildungslehre*. Tübingen: Niemeyer
- Ramers, Karl-Heinz & Heinz Vater. 1992. *Einführung in die Phonologie*. Hürth: Gabel-Verlag.
- Reis, Marga. 1983. Gegen die Kompositionstheorie der Affigierung. *Zeitschrift für Sprachwissenschaft* 2. 110-29.
- Reynolds, Bill. 1994. *Variation and phonological theory*. Philadelphia, PA: University of Pennsylvania dissertation.
- Rice, Curt. 2004. Optimizing Russian Gender. Handout for a talk presented at Formal Approaches to Slavic Linguistics. University of South Carolina, February 28, 2004.
- Rice, Curt. 2005. Optimizing Russian gender: A preliminary analysis. In Steven Franks, Frank Y. Gladney & Mila Tasseva-Kurktchieva (eds), *Formal Approaches to Slavic Linguistics 13: The South Carolina Meeting*, 265-275. Ann Arbor, MI: Michigan Slavic Publications.
- Rice, Curt. 2006. Optimizing gender. *Lingua* 116. 1394-1417.
- Rice, Sally. 1996. Prepositional prototypes. In Martin Pütz & René Dirven (eds), *The construal space in language and thought*, 135-165. Berlin: de Gruyter.
- Riemer, Nick. 2010. *Introducing Semantics*. Cambridge: Cambridge University Press.

- Robins, Robert H. 1997. *A Short History of Linguistics*, 4th edn. London/New York: Longman.
- Rodina Yulia & Marit Westergaard. 2013. Two gender systems in one mind. The acquisition of grammatical gender in Norwegian-Russian bilinguals. In Peter Siemund, Ingrid Gogolin, Monika Edith Schulz & Julia Davydova (eds), *Multilingualism and language diversity in urban areas: Acquisition, identities, space, education*, 95-126. Amsterdam: John Benjamins.
- Rogers, Margaret. 1987. Learners' difficulties with grammatical gender in German as a foreign language. *Applied Linguistics* 8(1). 48-74.
- Ronneberger-Sibold, Elke. 2004. Deutsch (Indogermanisch: Germanisch). In Geert E. Booij, Christian Lehmann & Joachim Mugdan (eds), *Morphologie: Ein Internationales Handbuch zur Flexion und Wortbildung*, Volume 2, 1267-1285. Berlin: de Gruyter.
- Rosch, Eleanor. 1973. Natural Categories. *Cognitive Psychology* 4. 328-50.
- Rosch, Eleanor. 1975a. Cognitive reference points. *Cognitive Psychology* 7. 532-34.
- Rosch, Eleanor. 1975b. Cognitive representations of semantic categories. *Journal of Experimental Psychology* 104(3). 192-233.
- Rosch, Eleanor. 1977. Human categorisation. In Neil Warren (ed.), *Advances in cross-cultural psychology*, 1-49. London: Academic Press.
- Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch & Barbara B. Lloyd (eds), *Cognition and categorization*. Hillsdale, NJ: Erlbaum.
- Rosch, Eleanor. 1998. Is wisdom in the brain? *Psychological Science* 10. 222-224.
- Rosch, Eleanor. 1999. Principles of categorization. In Eric Margolis & Stephen Laurence (eds), *Concepts: core readings*, 189-206. Cambridge, MA: MIT Press.
- Rosch, Eleanor & Carolyn Mervis. 1975. Family resemblances: studies in the internal structure of categories. *Cognitive Psychology* 7. 573-605.
- Rosch, Eleanor, Carolyn Mervis, Wayne Gray, David Johnson & Penny Boyes-Braem. 1976. Basic objects in natural categories. *Cognitive Psychology* 8. 382-439.
- Rosengren, Inger. 1977. *Ein Frequenzwörterbuch der deutschen Zeitungssprache*, Volume 2. Lund: Gleerup.
- Rosenbach, Annette. 2002. *Genitive variation in English*. Berlin: de Gruyter.
- Rosenthal, Saul H. 2009. *The rules for the gender of French nouns*, 4th edn. Arizona: Wheatmark.
- Ruhl, Charles. 1989. *On monosemy: A study in linguistic semantics*. Albany, New York: State University of New York Press.
- Rumshisky, Anna & Amber Stubbs 2017. Machine learning for higher-level linguistic tasks. In Nancy Ide & James Pustejovsky (eds), *Handbook of linguistic annotation*, 333-352. Dordrecht: Springer.
- Ruoff, Arno. 1981. *Häufigkeitswörterbuch gesprochener Sprache*. Tübingen: Niemeyer.

- Russ, Charles. 1994. *The German language today: A linguistic introduction*. London/New York: Routledge.
- Sadek, Carmen Sánchez, Jackie Kiraithe & Hildebrando Villarreal. 1975. The acquisition of the concept of grammatical gender in monolingual and bilingual speakers of Spanish. In Frances Aid, Melvin Resnick & Bohdan Saciuk (eds), *Colloquium on Hispanic linguistics*, 131-149. Washington: Georgetown University Press.
- Salmons, Joseph. 1993. The structure of the lexicon: Evidence from German gender assignment. *Studies in Language* 17(2). 411-435.
- Salmons, Joseph. 1994. Sketch of an interlanguage rule system: Advanced non-native German gender assignment. In Carol A. Blackshire-Belay (ed.), *Current issues in second language acquisition and development*, 187-204. Maryland/London: University Press of America.
- Scalise, Sergio. 1988. The notion of 'head' in morphology. In Geert E. Booij & Jaap van Marle (eds), *Yearbook of morphology 1988*, 229-245. Dordrecht: Kluwer.
- Scalise, Sergio. 1992. Compounding in Italian. *Rivista di Linguistica* 4. 175-199.
- Scalise, Sergio, Antonio Fábregas & Francesca Forza. 2009. Exocentricity in compounding. *Gengo Kenkyu* 135. 49-84.
- Scalise, Sergio & Antonio Fábregas. 2010. The head in compounding. In Sergio Scalise & Irene Vogel (eds), *Cross-disciplinary issues in compounding*, 109-126. Amsterdam: John Benjamins.
- Scheibl, György. 2008. Genusparameter in der Diskussion. *Zeitschrift für germanistische Linguistik* 36(1). 48-73.
- Schiller, Niels O. & Alfonso Caramazza. 2003. Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language* 48(1). 169-194.
- Schiller Niels O., Thomas Muentz, Iemke Horemans & Bernadette M. Jansma. 2003. The influence of semantic and phonological factors on syntactic decisions: An event-related brain potential study. *Psychophysiology* 40. 869-877.
- Schirmer, Lars. 2013. Regelgeleitete Genusvermittlung bei DaF-Lernenden in Mexiko. In Christian Fandrych, Adriana R. Galván Torres, Werner Heidermann, Ulrike Pleß & Erwin Tschirner (eds), *Text, Diskurs und Translation im Wandel*, 85-100. Tübingen: Stauffenberg.
- Schlick, Werner. 1984. Die Kriterien für die deutsche Genuszuweisung bei substantivischen Anglizismen. *The German Quarterly* 57(3). 402-431.
- Schlücker, Barbara & Matthias Hüning. 2009. Compounds and phrases. A functional comparison between German A+N compounds and corresponding phrases. *Rivista di Linguistica* 21(1). 209-234.
- Schlücker, Barbara & Ingo Plag. 2011. Compound or phrase? Analogy in naming. *Lingua* 121. 1539-1551.
- Schmid, Monika. 2002. *First language attrition, use and maintenance: The case of German Jews in anglophone countries*. Amsterdam: John Benjamins.

- Schmid, Tanja. 2001. OT accounts of optionality: Global ties and neutralization – a comparison. In Gereon Müller & Wolfgang Sternefeld (eds), *Competition in syntax*, 283-319. Berlin: de Gruyter.
- Schuler, Kathryn, Charles Yang & Elissa Newport. 2016. Testing the Tolerance Principle: Children form productive rules when it is more computationally efficient to do so. *The 38th Cognitive Science Society Annual Meeting*. 2321-2326.
- Schulte-Beckhausen, Marion. 2002. *Genusschwankung bei englischen, französischen, italienischen und spanischen Lehnwörtern im Deutschen. Eine Untersuchung auf Grundlage deutscher Wörterbücher seit 1945*. Frankfurt am Main: Peter Lang.
- Schwarze, Brigitte. 2008. *Genus im Sprachvergleich: Klassifikation und Kongruenz im Französischen, Spanischen und Deutschen*. Tübingen: Narr.
- Schwichtenberg, Beate & Niels Schiller. 2004. Semantic gender assignment regularities in German. *Brain and Language* 90. 326-337.
- Selkirk, Elisabeth. 1982. *The syntax of words*. Cambridge, MA: MIT Press.
- Semenza, Carlo & Sara Mondini. 2006. Compound words in neuropsychology. *Linguistische Berichte* 17. 331-348.
- Semplicini, Chiara. 2012. Dutch double gender nouns: Arbitrary or motivated agreement? *Journal of Germanic Linguistics* 24(2). 133-186.
- Simon, Horst J. & Heike Wiese. 2011. What are exceptions? And what can be done about them? In Horst J. Simon & Heike Wiese (eds), *Expecting the unexpected: Exceptions in grammar*, 3-30. Berlin: de Gruyter.
- Slobin, Dan. 1973. Cognitive prerequisites for the development of grammar. In Charles Ferguson & Dan Slobin (eds), *Studies of child language development*, 175-208. New York: Holt, Rinehart & Winston.
- Smith, Noah A. 2011. *Linguistic structure prediction*. San Rafael, CA: Morgan & Claypool.
- Sonnenstuhl, Ingrid, Sonja Eisenbeiss & Harald Clahsen. 1999. Morphological priming in the German mental lexicon. *Cognition* 72. 203-236.
- Spitz, Erich. 1965. Beitrag zur Genusbestimmung der deutschen Substantive. *Deutsch als Fremdsprache* 2(4). 35-43.
- Steen, Gerard. 2007. *Finding metaphor in grammar and usage*. Amsterdam: John Benjamins.
- Steinmetz, Donald. 1985. Gender in German and Icelandic: inanimate nouns. In Jan T. Faarlund (ed.), *Germanic Linguistics: Papers from a Symposium at the University of Chicago*, 10-28. Bloomington: Indiana University Linguistics Club.
- Steinmetz, Donald. 1986. Two principles and some rules for gender in German: inanimate nouns. *Word* 37. 189-217.

- Steinmetz, Donald. 2001. The Great Gender Shift and the attrition of neuter nouns in West Germanic: the example of German. In Irmengard Rauch & Gerald Carr (eds), *New Insights in Germanic Linguistics*, 201-223. Frankfurt am Main: Peter Lang.
- Steinmetz, Donald. 2006. Gender shifts in Germanic and Slavic: semantic motivation for neuter? *Lingua* 116. 1418-1440.
- Steinmetz, Donald & Curt Rice. 1989. The gender of inanimate nouns in German and Dutch. *Texas Linguistic Forum* 31. 157-189.
- Štekauer, Pavol. 1996. *A Theory of Conversion in English*. Frankfurt am Main: Peter Lang.
- Strandskogen, Åse-Berit & Rolf Strandskogen. 1986. *Norwegian: An Essential Grammar*. London: Routledge.
- Superanskaja, A. V. 1965. Rod zaimstvovannyx suščestivel'nyx v sovrenennom russkom jazyke [The gender of borrowed nouns in modern Russian]. *Voprosy kul'tury rec'ì* 6: 44-58.
- Szagun, Gisela, Barbara Stumper, Nina Sondag & Melanie Frank. 2007. The acquisition of gender marking by young German-speaking children: Evidence for learning guided by phonological regularities. *Journal of Child Language* 34. 445-471.
- Taft, Marcus & Fanny Meunier. 1998. Lexical representation of gender: A quasiregular domain. *Journal of Psycholinguistic Research* 27. 23-45.
- Talanga, Tomislav. 1987. *Das Phänomen der Genusschwankung in der deutschen Gegenwartssprache – untersucht nach Angaben neuerer Wörterbücher der deutschen Standardsprache*. Bonn: Rheinische Friedrich-Wilhelms-Universität Bonn dissertation.
- Taylor, John R. 1995. *Linguistic categorization: Prototypes in linguistic theory*. Oxford: Clarendon Press.
- Taylor, John R. 2003. *Linguistic categorization*, 3rd edn. Oxford: Oxford University Press.
- Thieroff, Rolf & Petra M. Vogel. 2012. *Flexion*, 2nd edn. Heidelberg: Universitätsverlag Winter.
- Thomoglou, Polichronia. 2004. *Genuserwerb bei griechischen Lernern des Deutschen*. Frankfurt am Main: Peter Lang.
- Thornton, Anna. 2001. Some reflections on gender and inflectional class assignment in Italian. In Chris Schaner-Wolles, John Rennison & Friedrich Neubarth (eds), *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler*, 479-87. Turin: Rosenberg & Sellier.
- Thornton, Anna. 2009. Constraining gender assignment rules. *Language Sciences* 31. 14-32.
- Toman, Jindřich. 2001. Word syntax. In Andrew Spencer & Arnold M. Zwicky (eds), *The handbook of morphology*, 306-321. Oxford: Blackwell.
- Topintzi, Nina. 2005a. Stress shift in Arabela: the effect of onsets. Talk presented at the 3rd Cambridge Postgraduate Conference in Language Research, 18 March 2005, Cambridge, UK.

- Topintzi, Nina. 2005b. Putting 'commas' at the right place: A note on crucial non-ranking in OT. *UCL Working Papers in Linguistics* 17. 83-108.
- Trosterud, Trond. 2001. Genustilordning i norsk er regelstyrt [Gender assignment in Norwegian is rule-governed]. *Norsk Lingvistisk Tidsskrift* 19. 29-58.
- Tucker, G. Richard, Wallace Earl Lambert, André Rigault & Norman Segalowitz. 1968. A psycholinguistic investigation of French speakers' skill with grammatical gender. *Journal of Verbal Learning and Verbal Behaviour* 7. 312-316.
- Tucker, G. Richard, Wallace Earl Lambert & André Rigault. 1977. *The French speaker's skill with grammatical gender: An example of rule-governed behaviour*. The Hague & Paris: Mouton.
- Tuggy, David. 1993. Ambiguity, polysemy, and vagueness. *Cognitive linguistics* 4(3). 273-290.
- Tuten, Donald, Lucia Caycedo Garner & Carmelo Esterrich. 2011. *Fuentes: Lectura y redacción*, 4th edn. Boston, MA: Heinle.
- Tversky, Barbara. 1985. Components and categorization. In Colette Craig (ed.), *Categorization and noun classification*, 63-76. Amsterdam: John Benjamins.
- Tyler, Andrea & Vyvyan Evans. 2001. Reconsidering prepositional polysemy networks: The case of over. *Language* 77(4). 724-765.
- Tyler, Andrea & Vyvyan Evans. 2003. *The semantics of English prepositions*. Cambridge: Cambridge University Press.
- Umbreit, Birgit. 2010. 'Does love come from to love or to love from love?' Why lexical motivation has to be regarded as bidirectional. In Alexander Onysko & Sascha Michel (eds), *Cognitive Perspectives on Word Formation*, 301-334. Berlin: de Gruyter.
- Umbreit, Birgit. 2015. *Zur Direktionalität der lexikalischen Motivation*. Tübingen: Narr.
- Urgelles-Coll, Miriam. 2010. *The syntax and semantics of discourse markers*. London/Oxford: Bloomsbury.
- Vigliocco, Gabriella, Tiziana Antonini & Merrill F. Garrett. 1997. Grammatical gender is on the tip of Italian tongues. *Psychological Science* 8, 314-317.
- Wahrig Deutsches Wörterbuch*, 6th edn. 1997. Mannheim: Brockhaus.
- Walter, S. 1975. *Zur Entwicklung morphologischer Strukturen bei Kindern*. Ms. Universität Heidelberg, Psychologisches Institut.
- Wängler, Hans-Heinrich. 1963. *Rangwörterbuch hochdeutscher Umgangssprache*. Marburg: Eiwert.
- Watts, Sheila. 2016. What didn't German do? *The 17th Diachronic Generative Syntax Conference, Ghent, June 2016*. Unpublished conference paper.
- Wegener, Heide. 1995. *Die Nominalflexion des Deutschen - verstanden als Lerngegenstand*. Tübingen: Niemeyer.

- Wegener, Heide. 2000. German gender in children's second language acquisition. In Barbara Unterbeck & Matti Rissanen (eds), *Gender in grammar and cognition*, 511-544. Berlin: de Gruyter.
- Wegera, Klaus-Peter. 1997. *Das Genus. Ein Beitrag zur Didaktik des DaF-Unterrichts*. Munich: Iudicium.
- Weinrich, Harald. 2007. *Textgrammatik der deutschen Sprache*, 4th edn. Hildesheim: Olms.
- Weißefeld, Frank. 2004. *Die Gestaltungsproblematik der Unternehmensnachfolge im personengeprägten Mittelstand*. Hamburg: Diplomica.
- Wiese, Richard. 1996. *The phonology of German*. Oxford: Clarendon Press.
- Wiesinger, Peter. 1983. Einleitung der deutschen Dialekte: Die Gliederung der deutschen Dialekte. In Werner Besch, Ulrich Knoop, Wolfgang Putschke & Herbert Wiegand (eds), *Dialektologie: Ein Handbuch zur deutschen und allgemeinen Dialektforschung*, 807-900. Berlin: de Gruyter.
- Wiktionary. 2013. Dutch words suffixed with *-uur*. https://en.wiktionary.org/wiki/Category:Dutch_words_suffixed_with_-uur. (20 June, 2018).
- Wiktionary. 2017a. *-eder*. <https://en.wiktionary.org/wiki/-eder>. (20 June, 2018).
- Wiktionary. 2017b. *-loog*. <https://en.wiktionary.org/wiki/loog#Dutch>. (20 June, 2018).
- Wiktionary. 2018a. *-ant*. <https://en.wiktionary.org/wiki/-ant#Dutch>. (20 June, 2018).
- Wiktionary. 2018b. *-tie*. <https://en.wiktionary.org/wiki/-tie#Dutch>. (20 June, 2018).
- Wiktionary. 2018c. *-ent*. https://en.wiktionary.org/wiki/-ent#Suffix_2. (20 June, 2018).
- Wiktionary. 2018d. *-teca*. <https://en.wiktionary.org/wiki/teca>. (20 June, 2018).
- Wilder, Chris & Hans-Martin Gärtner. 1997. Introduction. In Chris Wilder & Hans-Martin Gärtner & Manfred Bierwisch (eds), *The role of economy principles in linguistic theory*, 1-35. Berlin: Akademie Verlag.
- Williams, Edwin. 1981. On the notions 'lexically related' and 'head of a word'. *Linguistic Inquiry* 12. 245-274.
- Williams, John N. 1992. Processing polysemous words in context: Evidence for interrelated meanings. *Journal of Psycholinguistic Research* 21(3). 193-218.
- Wittgenstein, Ludwig. 1953. *Philosophical investigations*. New York: Macmillan.
- Wouden, Ton van der. 2016a. *-er* (nominal). *Taalportaal*. www.taalportaal.org/taalportaal/topic/pid/topic-14093156822693140. (6 September, 2016).
- Wouden, Ton van der. 2016b. *-iaan, -aan*. *Taalportaal*. www.taalportaal.org/taalportaal/topic/pid/topic-14243602175219657. (6 September, 2016).
- Wouden, Ton van der. 2016c. *-ist*. *Taalportaal*. www.taalportaal.org/taalportaal/topic/pid/topic-14332576179168771 (6 September, 2016).

- Wouden, Ton van der. 2017. Chemical morphology. *Taalportaal*. www.taalportaal.org/taalportaal/topic/pid/topic-14011042154732730. (10 August, 2017).
- Wurzel, Wolfgang Ullrich. 1984. *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie-Verlag.
- Wurzel, Wolfgang Ullrich. 1994. Gibt es im Deutschen noch eine einheitliche Substantivflexion? oder: Auf welche Weise ist die deutsche Substantivflexion möglichst angemessen zu erfassen? In Klaus-Michael Köpcke (ed.), *Funktionale Untersuchungen zur deutschen Nominal- und Verbalmorphologie*, 29-44. Tübingen: Narr.
- Yang, Charles. 2005. On productivity. In Pierre Pica, Johan Rooryck & Jeroen van Craenenbroeck (eds), *Linguistic variation yearbook 5*, 265-302. Amsterdam: John Benjamins.
- Yang, Wenliang. 1990. *Anglizismen im Deutschen am Beispiel des Nachrichtenmagazins ‚Der Spiegel‘*. Berlin: de Gruyter.
- Zekhnini, A. & Jan H. Hulstijn. 1995. An experimental study on the learning of arbitrary and nonarbitrary gender of pseudo Dutch nouns by nonnative and native speakers of Dutch. *Toegepaste Taalwetenschap in Artikelen* 53. 121-136.
- Zifonun, Gisela, Ludger Hoffman, Bruno Strecker & Joachim Ballweg. 1997. *Grammatik der deutschen Sprache*. Berlin: de Gruyter.
- Zipf, George K. 1935. *The psychobiology of language*. Houghton-Mifflin.
- Zubin, David & Klaus-Michael Köpcke. 1981. Gender: A less than arbitrary grammatical category. *Chicago Linguistic Society* 17. 439-449.
- Zubin, David & Klaus-Michael Köpcke. 1984. Affect classification in the German gender system. *Lingua* 63(1). 41-96.
- Zubin, David & Klaus-Michael Köpcke. 1986. Gender and folk taxonomy. The indexical relation between grammatical and lexical categorization. In Colette Craig (ed.), *Noun classification and categorization*, 139-180. Amsterdam: John Benjamins.
- Zwanenburg, Wiecher. 1983. *Productivité morphologique et emprunt* [Morphological productivity and borrowing]. Amsterdam: John Benjamins.
- Zwanenburg, Wiecher. 1992. Compounding in French. *Rivista di Linguistica* 4(1). 221-239.
- Zwicky, Arnold M. 1985. Heads. *Journal of Linguistics* 21(1). 1-29.

Appendices

Appendix A: The sample

<i>das Abgangszeugnis</i>	<i>der Dipteros</i>	<i>der Tausender</i>	<i>die Konnivenz</i>
<i>das Allegro</i>	<i>der Docker</i>	<i>der Test</i>	<i>die Konvergenz</i>
<i>das Anachoretentum</i>	<i>der Dom</i>	<i>der Thyristor</i>	<i>die Korpuskulartheorie</i>
<i>das Anderkonto</i>	<i>der Dorfclub, Dorfklub</i>	<i>der Topverdiener</i>	<i>die Krankenanstalt</i>
<i>das Antependium</i>	<i>der Drall</i>	<i>der Tourenski, Tourensch</i>	<i>die Kriegsverbrecherin</i>
<i>das Arrival</i>	<i>der Dreiklang</i>	<i>der Tran</i>	<i>die Kunst</i>
<i>das Bakschisch</i>	<i>der Dual</i>	<i>der Umformer</i>	<i>die Kursbewegung</i>
<i>das Benediktenkraut</i>	<i>der Dweil</i>	<i>der Undank</i>	<i>die Lafette</i>
<i>das Betäubungsmittel</i>	<i>der Ehestreit</i>	<i>der Ungeist</i>	<i>die Laryngitis</i>
<i>das Blackjack, Black Jack</i>	<i>der Einriss</i>	<i>der Unsegen</i>	<i>die Legalität</i>
<i>das Bootcamp</i>	<i>der Einstundentakt</i>	<i>der Unterführer</i>	<i>die Leiste</i>
<i>das Coaching</i>	<i>der Eisenbahner</i>	<i>der Vegetarier</i>	<i>die Lese</i>
<i>das Dazutun</i>	<i>der Encoder, Enkoder</i>	<i>der Verbesserer</i>	<i>die Liasformation</i>
<i>das Deleatur</i>	<i>der Eremit</i>	<i>der Verdruss</i>	<i>die Liederlichkeit</i>
<i>das Derivat</i>	<i>der Erweiterungsbau</i>	<i>der Verkauf</i>	<i>die Lyssa</i>
<i>das Deutschsprechen</i>	<i>der Esslöffel</i>	<i>der Vielflieger</i>	<i>die Magenschleimhaut</i>
<i>das Diaphanbild</i>	<i>der Eunuch</i>	<i>der Vogelschutz</i>	<i>die Malaria</i>
<i>das Drogengeschäft</i>	<i>der Falschspieler</i>	<i>der Vollender</i>	<i>die Marke</i>
<i>das Eclair</i>	<i>der Federling</i>	<i>der Voraus</i>	<i>die MAZ (magnetische Bildaufzeichnung)</i>
<i>das Eigentumsrecht</i>	<i>der Feldverweis</i>	<i>der Wartestand</i>	<i>die Messerspitze</i>
<i>das Epistyl</i>	<i>der Fetischist</i>	<i>der Webstuhl</i>	<i>die Mimin</i>
<i>das Erschöpfungssyndrom</i>	<i>der Fladerschnitt</i>	<i>der Wehrpass</i>	<i>die Modallogik</i>
<i>das Fahrenheit</i>	<i>der Fluchtversuch</i>	<i>der Weitschuss</i>	<i>die Mühsal</i>
<i>das Farsi</i>	<i>der Folklorist</i>	<i>der Werkbus</i>	<i>die Munterkeit</i>
<i>das Ferrum</i>	<i>der Fraktionssprecher</i>	<i>der Wettbewerbsnachteil</i>	<i>die Nachäfferei</i>
<i>das Filet</i>	<i>der Fremdstoff</i>	<i>der Wiedehopf</i>	<i>die Neotenie</i>
<i>das Flettnerruder</i>	<i>der F-Schlüssel</i>	<i>der Windhund</i>	<i>die Neubesetzung</i>
<i>das Freixemplar</i>	<i>der Galvaniseur</i>	<i>der Zirkelkasten</i>	<i>die Niederdruckheizung</i>
<i>das Fungizid</i>	<i>der Geograf, Geograph</i>	<i>der Zusammenbruch</i>	<i>die Nikotinhaltigkeit</i>
<i>das Gebärdenspiel</i>	<i>der Geschäftsschluss</i>	<i>der Zweiteiler</i>	<i>die Noxe</i>
<i>das Gegenüber</i>	<i>der Glockenmantel</i>	<i>die ABM (Arbeitsbeschaffungs- maßnahme)</i>	<i>die Oberin</i>
<i>das Geleucht</i>	<i>der Gottmensch</i>	<i>die Abseitsposition</i>	<i>die Occasion, Okkasion</i>
<i>das Getändel</i>	<i>der Großbuchstabe</i>	<i>die Abwärme</i>	<i>die ÖNB (Österreichische Nationalbibliothek)</i>
<i>das Goldhähnchen</i>	<i>der Grufti</i>	<i>die Adduktion</i>	<i>die Optionsanleihe</i>
<i>das Haben</i>	<i>der Handball</i>	<i>die Agrogentechnik</i>	<i>die Orgelpfeife</i>
<i>das Hallenbad</i>	<i>der Hannoveraner</i>	<i>die Akrozephalie</i>	<i>die Pädagogisierung</i>
<i>das Hexaeder</i>	<i>der Hasenpfeffer</i>	<i>die Alchimie</i>	<i>die Palpe</i>
<i>das His</i>	<i>der Hausmeister</i>	<i>die All-Star-Band</i>	<i>die Panzerdivision</i>
<i>das Hochzeitsgeschenk</i>	<i>der Heißhunger</i>	<i>die Alternanz</i>	<i>die Paralalie</i>
<i>das Indossament</i>	<i>der Helvetier</i>	<i>die Amelioration</i>	<i>die Parkkralle</i>
<i>das Informel</i>	<i>der Herumtreiber</i>	<i>die Anknüpfung</i>	<i>die Pauschalität</i>
<i>das Inkrafttreten</i>	<i>der Himmelskörper</i>	<i>die Antwort</i>	<i>die Pelzware</i>
<i>das J <Buchstabe></i>	<i>der Hochgenuss</i>	<i>die Apokope</i>	<i>die Perfidie</i>
<i>das Juchtenleder</i>	<i>der Hops</i>	<i>die Aszetik</i>	<i>die Phänomenologie</i>
<i>das Jüngstenrecht</i>	<i>der Hunderter</i>	<i>die Aufgliederung</i>	<i>die Pimpernuss</i>
<i>das Kilovolt</i>	<i>der Hypothekenzins</i>	<i>die Aufsichtung</i>	<i>die Plagiatorin</i>
<i>das Kontaktformular</i>	<i>der Intendant</i>	<i>die Ausreifung</i>	<i>die Pluralisierung</i>
<i>das Kopfzerbrechen</i>	<i>der Interrogativsatz</i>	<i>die Äußerung</i>	<i>die Polymerisierung</i>
<i>das Langhaus</i>	<i>der Islam</i>	<i>die Autarkie</i>	<i>die Präsentation</i>
<i>das Leben</i>	<i>der Jetztmensch</i>	<i>die Bandage</i>	<i>die Priorin</i>
<i>das Liniment</i>	<i>der Kalamit</i>	<i>die Bärentraube</i>	<i>die Produktlinie</i>
<i>das Lormalphabet</i>	<i>der Kapuzineraffe</i>	<i>die Bassgeige</i>	<i>die Prospektierung</i>
<i>das Lunarium</i>	<i>der Karlsbader</i>	<i>die Begriffsbestimmung</i>	<i>die Prozessordnung</i>
<i>das Maßhalten</i>	<i>der Kauerstart</i>	<i>die Bekohlung</i>	<i>die Pupille</i>
<i>das Meerschweinchen</i>	<i>der Key-Account-Manager</i>	<i>die Bessemerbirne</i>	<i>die Quarte</i>
<i>das Mikrofarad</i>	<i>der Kirschegeist</i>	<i>die Bildhauerkunst</i>	<i>die Quetschfalte</i>
<i>das Miteinander</i>	<i>der Knacker</i>	<i>die Biomedizin</i>	
<i>das Monomer</i>	<i>der Kohlenherd</i>		

<i>das Motiv</i>	<i>der Konformismus</i>	<i>die Blausäure</i>	<i>die Rackerei</i>
<i>das Nomen proprium</i>	<i>der Koton</i>	<i>die Bloßstellung</i>	<i>die Ratlosigkeit</i>
<i>das Normalmaß</i>	<i>der Kreisdurchmesser</i>	<i>die Bodybuilderin</i>	<i>die Rechtswissenschaft</i>
<i>das Nuteisen</i>	<i>der Krumme</i>	<i>die Bougainvillea</i>	<i>die Regel</i>
<i>das Oktavformat</i>	<i>der Kühlwagen</i>	<i>die Bremsrakete</i>	<i>die Regung</i>
<i>das Osmium</i>	<i>der Landadel</i>	<i>die Buchführung</i>	<i>die Renitenz</i>
<i>das Partizip</i>	<i>der Läufer</i>	<i>die Cantate, Kantate</i>	<i>die Reseda</i>
<i>das Perpetuum mobile</i>	<i>der Leichenwagen</i>	<i>die C-Dur-Tonleiter</i>	<i>die Retourkutsche</i>
<i>das Polarisationsmikroskop</i>	<i>der Literaturnobelpreis</i>	<i>die Chargennummer</i>	<i>die Rezeptur</i>
<i>das Postgiroamt</i>	<i>der Löffelstiel</i>	<i>die Chromatografie</i>	<i>die Richtung</i>
<i>das Q <Buchstabe></i>	<i>der Lorenzstrom</i>	<i>die Cranberry</i>	<i>die Rinne</i>
<i>das Rauschgold</i>	<i>der Löwenbändiger</i>	<i>die Dahlie</i>	<i>die Rockoper</i>
<i>das Rektapapier</i>	<i>der Mainzer</i>	<i>die Daphnia, Daphnie</i>	<i>die Rupiah</i>
<i>das Rondeau</i>	<i>der Mandelbaum</i>	<i>die Dasselbeule</i>	<i>die Sachlichkeit</i>
<i>das Rubidium</i>	<i>der Manteltarif</i>	<i>die Dendrochronologie</i>	<i>die Salatkartoffel</i>
<i>das Säumchen</i>	<i>der Martinstag</i>	<i>die Detailkenntnis</i>	<i>die Sanitärtechnik</i>
<i>das Schaffen</i>	<i>der Matsch-und-Schnee-Reifen</i>	<i>die Disproportion</i>	<i>die Schaufensterpuppe</i>
<i>das Schembartlaufen</i>	<i>der Meisterdetektiv</i>	<i>die Dunstglocke</i>	<i>die Schwachstelle</i>
<i>das Schienbein</i>	<i>der Menschegeist</i>	<i>die Durchhalteparole</i>	<i>die Schwarzarbeit</i>
<i>das Schlittschuhlaufen</i>	<i>der Metropolit</i>	<i>die Durchschrift</i>	<i>die Schweißperle</i>
<i>das Schnällchen</i>	<i>der Mischer</i>	<i>die Einehe</i>	<i>die Schwestertracht</i>
<i>das Schnippchen</i>	<i>der Mittelstürmer</i>	<i>die Einigkeit</i>	<i>die Sehnenzerrung</i>
<i>das Sophisma</i>	<i>der Molekularbiologe</i>	<i>die Elektrizität</i>	<i>die Selbstständigkeit</i>
<i>das Spektroskop</i>	<i>der Moralprediger</i>	<i>die EM (Europameisterschaft)</i>	<i>die Seniorengymnastik</i>
<i>das Stop-and-go</i>	<i>der Mustang</i>	<i>die Enkelin</i>	<i>die Serologie</i>
<i>das Supremum</i>	<i>der Nachschlag</i>	<i>die Entkriminalisierung</i>	<i>die Shapingmaschine</i>
<i>das Tamtam</i>	<i>der Nagel</i>	<i>die Entvölkerung</i>	<i>die Sowjetrepublik</i>
<i>das Tempo</i>	<i>der Nanga Parbat</i>	<i>die Erkundigung</i>	<i>die Spielpause</i>
<i>das Tipi</i>	<i>der Nationalstaat</i>	<i>die Examensangst</i>	<i>die Sprachgrenze</i>
<i>das Triolett</i>	<i>der Nebenbuhler</i>	<i>die Expertengruppe</i>	<i>die Spundbohle</i>
<i>das Uhrband</i>	<i>der New-Orleans-Jazz</i>	<i>die Exzellenzinitiative</i>	<i>die Stadtentwicklung</i>
<i>das Universalmittel</i>	<i>der Offiziersanwärter</i>	<i>die Finsternis</i>	<i>die Steelband</i>
<i>das UhrG</i>	<i>der Out-of-area-Einsatz</i>	<i>die Fortifikation</i>	<i>die Stilllegung, Still-Legung</i>
<i>(Urheberrechtsgesetz)</i>	<i>der Patagonier</i>	<i>die Frittenbude</i>	<i>die Strenge</i>
<i>das Vadium</i>	<i>der Petent</i>	<i>die Fustage</i>	<i>die Strukturtapeete</i>
<i>das Vorstadium</i>	<i>der Pfifferling</i>	<i>die Garonne</i>	<i>die Stundenzahl</i>
<i>das Waltharilied</i>	<i>der Phönizier</i>	<i>die Gedunsenheit</i>	<i>die Substandardwohnung</i>
<i>das Weißbuch</i>	<i>der Pickhammer</i>	<i>die Geisel</i>	<i>die Synärese</i>
<i>das Wissen</i>	<i>der Platzanweiser</i>	<i>die Generalagentin</i>	<i>die Szientistin</i>
<i>das Wohlverhalten</i>	<i>der Porphyrt</i>	<i>die Gerichtsmedizin</i>	<i>die Tagereise</i>
<i>das ZDF (Zweites Deutsches Fernsehen)</i>	<i>der Präfekt</i>	<i>die Gesetzgebung</i>	<i>die Taschenausgabe</i>
<i>das Zeitungslesen</i>	<i>der Presenter</i>	<i>die Gewinnquote</i>	<i>die Teilbarkeit</i>
<i>das Zusammenwirken</i>	<i>der Prolaps</i>	<i>die Gimpe</i>	<i>die Thematik</i>
<i>das Zyan, Cyan</i>	<i>der Publikums geschmack</i>	<i>die Glazialfauna</i>	<i>die Toilette</i>
<i>der Abend</i>	<i>der Rähm</i>	<i>die Grandel</i>	<i>die Trassenführung</i>
<i>der Achter</i>	<i>der Rapfen</i>	<i>die Gräzität</i>	<i>die Tretmine</i>
<i>der Affenbrotbaum</i>	<i>der Rechercheauftrag</i>	<i>die GST (Gesellschaft für Sport und Technik)</i>	<i>die Trope</i>
<i>der Angelpunkt</i>	<i>der Reinfall</i>	<i>die Gutgläubigkeit</i>	<i>die Tschechin</i>
<i>der Anreiz</i>	<i>der Rotang</i>	<i>die Hairstylistin</i>	<i>die Turbine</i>
<i>der Ara</i>	<i>der Rufer</i>	<i>die Hegemonie</i>	<i>die Tyrierin</i>
<i>der Areopag</i>	<i>der Samländer</i>	<i>die Herbst-Tagundnachtgleiche</i>	<i>die Übereile</i>
<i>der Asphalt</i>	<i>der Satellit</i>	<i>die Hintereinanderschaltung</i>	<i>die Überlegenheit</i>
<i>der Audiostream</i>	<i>der Schanker</i>	<i>die Holländerin</i>	<i>die Überstunde</i>
<i>der Aufzugsschacht, Aufzugsschacht</i>	<i>der Schiri (Schiedsrichter)</i>	<i>die Hubbrücke</i>	<i>die Umschau</i>
<i>der Ausdrusch</i>	<i>der Schlangenbeschwörer</i>	<i>die Hybris</i>	<i>die Umwicklung</i>
<i>der Ausheber</i>	<i>der Schlemmer</i>	<i>die Idiotin</i>	<i>die Unterschreitung</i>
<i>der Avistawechsel</i>	<i>der Schönfärber</i>	<i>die Imbezilität</i>	<i>die Unversehrtheit</i>
<i>der Backenbart</i>	<i>der Schrankkoffer</i>	<i>die Inadäquatheit</i>	<i>die Vergessenheit</i>
<i>der Bautzener</i>	<i>der Schrumpfergermane</i>	<i>die Insertion</i>	<i>die Verleitung</i>
<i>der Bedrücker</i>	<i>der Schund</i>	<i>die Inzahlungnahme</i>	<i>die Verpflegung</i>
<i>der Beigeschmack</i>	<i>der Sechsling</i>	<i>die Janitscharenmusik</i>	<i>die Verschlossenheit</i>
	<i>der Sekans</i>		<i>die Verspottung</i>
			<i>die Vertretungsvollmacht</i>

der Bergführer
der Besatz
der Beutenhonig
der Bezwingler
der Brandy
der Bronzit
der Bummelant
der Bürgersteig
der Byzantiner
der Chickenwing
der Containerverkehr
der DAAD
der Defroster
der Dienstgebrauch
der Dime

der Siderit
der Sikahirsch
der Sinnenrausch
der Skatabend
der Slowake
der Sog
der Sommer
der Sparerfreibetrag
der Spitzkühler
der Standartenträger
der Statist
der Stelzvögel
der Steuerfreibetrag
der Strampelanzug
der Suizident

die Kabinettskrise
die Kaltblütigkeit
die Kampfmaßnahme
die Kantonalbank
die Karwoche
die Kasusendung
die Kelle
die Kimm
die Klasse
die Kleinschreibung
die Klinge
die Knopflochseide
die Kolmation
die Kommentatorin
die Komplexbrigade

die Verzögerung
die Vindikation
die Vorherbestimmung
die Waagrechte
die Wahlverwandschaft
die Weltmarke
die Wienerin
die Wortkargheit
die Wutrede
die Zackenkrone
die Zergliederung
die Ziellinie
die Zoohandlung
die Zufälligkeit
die Zwangsschiene

Subset of 55 nouns

Antwerpen
Bautzen
Bessarabien
das Außerstreitverfahren
das Aviso
das Flett
das Kulm
das/der Ar
das/die Korpuskel
der Außerstreitrichter
der Aviso
der Detailhandel
der Farre
der Holländer₁

der Holländer₂
der Holländer₃
der Kolm (der/das Kulm)
der Rahm
der Sigrist
der Tip
der/das Avis
der/das Garnknäuel
der/das Juchten
der/das New Look
der/das Renforcé
die Angelobung
die Elektrische
die Metropolis

die Tramway
Frutti
Geisa (das Geison)
Gerichtskosten
Großbritannien
Iskariot
Karlovy Vary
Karlsbad
Klimt
Korpuskularstrahlen (der Strahl)
Kranke
Kühlungsborn
Lärtes
Münster

Nadja
Naftali
Nagasaki
New Hampshire
Patagonien
Rahel
Renitente
Schiele
Sechsfache
Sigrun
Sophia, Sophie
Tintoretto
Walter

Appendix B: The constraints

Constraint	Explanation	Source	Examples from sample
-(C)(C)/j/ MONO → * _F , * _N	Monosyllabic nouns with a [j] in final position cannot be feminine or neuter. The [j] may be preceded by up to two consonants. E.g. <i>der Klatsch</i> . Köpcke (1982: 102) includes all nouns ending in [tʃ] in this category, in spite of the fact that /tʃ/ is argued by some to be a single phoneme (see e.g. Ramers & Vater 1992: 85-91).	Altmann & Raettig (1973: 302), Köpcke (1982: 102)	<i>Jetztmensch, Gottmensch</i>
-(ER)EI → * _M , * _N	Nouns with the suffix <i>-(er)ei</i> cannot be masculine or neuter, e.g. <i>die Bäckerei, die Fischerei</i>	Mills (1986: 30), Flämig (1991: 453), Zifonun et al. (1997: 33), Hoberg (2004: 87), Thomoglou (2004: 41), Weinrich (2007: 327), Duden Grammatik (2009: 164), Engel (2009: 280), Eisenberg (2013: 133)	<i>Nachäfferei, Rackerei</i>
-(O)LOGE → * _F , * _N	Nouns with the (foreign) suffix <i>-(o)loge</i> cannot be feminine or neuter, e.g. <i>der Neurologe, Archäologe</i>	Altmann & Raettig (1973: 302), Flämig (1991: 453), Engel (2009: 280)	<i>Molekularbiologe</i>
-/r/ STOP (C) MONO → * _F , * _N	Monosyllabic nouns ending in /r/ followed by a stop followed by an optional consonant cannot be feminine or neuter, e.g. <i>der Hort</i>	Köpcke (1982: 101)	<i>Kohlenherd</i>
-/tʃ/;/Ntʃ/;/Nj/ MONO → * _F , * _N	Monosyllables ending in <i>-/tʃ/</i> , <i>-/Ntʃ/</i> or <i>-/Nj/</i> cannot be feminine or neuter, e.g. <i>der Klatsch</i>	Salmons (1993: 425)	<i>Gottmensch, Jetztmensch</i>
-[l]C MONO → * _F , * _N	Monosyllabic nouns whose coda consists of an [l] followed by a consonant cannot be feminine or neuter, e.g. <i>der Fels</i>	Köpcke (1982: 100)	<i>Kolm, Kilovolt, Rauschgold</i>
[j]C- MONO → * _F , * _N	Monosyllabic nouns with an onset consisting of [j] followed by a consonant cannot be feminine or neuter, e.g. <i>der Schrott</i>	Köpcke (1982: 89f.), Mills (1986: 33), Köpcke & Zubin (1997), Hoberg (2004: 98), Menzel (2004: 68), Chan (2005: 296)	<i>Löffelstiel, Kauerstart</i>
-A → * _M , * _N	Nouns ending in <i>-a</i> cannot be masculine or neuter, e.g. <i>die Algebra</i>	Steinmetz (1986: 197), Steinmetz & Rice (1989: 166), Thomoglou (2004: 41), Duden Grammatik (2009: 164), Durrell (2011: 8), Kraiss (2014: 36)	<i>Lyssa, Malaria</i>
-ADE → * _M , * _N	Nouns with the “foreign” suffix <i>-ade</i> cannot be masculine or neuter, e.g. <i>die Blockade</i>	Hoberg (2004: 87), Thomoglou (2004: 41), Weinrich (2007: 327), Duden Grammatik (2009: 164)	<i>Komplexbrigade</i>
-AGE → * _M , * _N	Nouns with the “foreign” suffix <i>-age</i> [a:ɜə] cannot be masculine or neuter, e.g. <i>die Massage</i>	Flämig (1991: 453), Hoberg (2004: 87), Thomoglou (2004: 41), Weinrich (2007: 327), Duden Grammatik (2009: 164), Engel (2009: 280)	<i>Fustage, Bandage</i>
-AL → * _M , * _F	Polysyllabic nouns ending in (stressed) [a:] <i>-al</i> cannot be masculine or feminine, e.g. <i>das Mineral</i>	Onysko (2007: 172), Bjornstad (2014: 78)	<i>Dual</i>
-AN → * _M , * _F	Polysyllabic nouns ending in <i>-an</i> cannot be masculine or feminine, e.g. <i>das Methan</i>	Flämig (1991: 453)	<i>Zyan</i>

-ANER→*F,*N	Nouns with the suffix <i>-aner</i> cannot be feminine or neuter, e.g. <i>der Chomskyaner</i>	Engel (2009: 279)	<i>Hannoveraner</i>
-ANG→*F,*N	Nouns ending in <i>-ang</i> cannot be feminine or neuter, e.g. <i>der Klang</i>	Altmann & Raettig (1973: 302)	<i>Rotang, Mustang, Dreiklang</i>
-ANT→*F,*N	Nouns with the “foreign” suffix <i>-ant</i> cannot be feminine or neuter, e.g. <i>der Protestant</i>	Flämig (1991: 453), Hoberg (2004: 86), Thomoglou (2004: 41), Weinrich (2007: 326), Duden Grammatik (2009: 164), Engel (2009: 279)	<i>Bummelant</i>
-ANZ→*M,*N	Nouns with the “foreign” suffix <i>-anz</i> cannot be masculine or neuter, e.g. <i>die Arroganz</i>	Hoberg (2004: 87), Menzel (2004: 66), Weinrich (2007: 327), Duden Grammatik (2009: 164), Engel (2009: 280)	<i>Alternanz</i>
-AR→*M,*F	Nouns ending in stressed <i>-ar</i> cannot be masculine or feminine, e.g. <i>das Honorar</i>	Nelson (1998: 216)	<i>Kontaktformular</i>
-AR SUFF.→*F,*N	Nouns with the “foreign” suffix <i>-ar</i> cannot be feminine or neuter, e.g. <i>der Bibliothekar</i>	Hoberg (2004: 89), Engel (2009: 280)	<i>Freiexemplar</i>
-AR SUFF.→*M,*F	Nouns with the “foreign” suffix <i>-ar</i> cannot be masculine or feminine, e.g. <i>das Inventar</i>	Hoberg (2004: 89), Durrell (2011: 10), Bjornstad (2014: 78)	
-AT→*M,*F	Nouns with the (stressed) “foreign” suffix <i>-at</i> (incl. <i>-iat</i>) cannot be masculine or feminine, e.g. <i>das Telefonat</i>	Flämig (1991: 453), Hoberg (2004: 89), Menzel (2004: 66), Weinrich (2007: 326), Engel (2009: 279)	<i>Derivat, Oktavformat</i>
-AT→*F,*N	Nouns with the (stressed) “foreign” suffix <i>-at</i> cannot be feminine or neuter, e.g. <i>der Stipendiat</i>	Hoberg (2004: 89), Engel (2009: 279)	
-B→*F	Nouns ending in <i>-b</i> cannot be feminine, e.g. <i>der Urlaub</i>	Altmann & Raettig (1973: 302)	<i>Dorfklub</i>
-C S MONO→*F,*N	Monosyllabic nouns whose coda consists of at least one consonant followed by /s/ cannot be feminine or neuter, e.g. <i>der Schutz</i> . Köpcke (1982: 102) includes all nouns ending in /ts/ and /ks/ in this category, in spite of the fact that /ts/ and /ks/ are argued by some to be single phonemes (see e.g. Ramers & Vater 1992: 85-91).	Altmann & Raettig (1973: 302), Köpcke (1982: 102)	<i>Hops, Interrogativsatz</i>
CCCVC MONO→*F,*N	Monosyllables with the structure CCCVC cannot be feminine or neuter, e.g. <i>der Sprung</i>	Köpcke (1982: 84)	<i>Lorenzstrom, Ehestreit</i>
CCVC MONO→*F,*N	Monosyllabic nouns with the structure CCVC cannot be feminine or neuter, e.g. <i>der Staat</i>	Köpcke (1982: 85)	<i>Drall, Tran</i>
CCVCC→*F,*N	Monosyllabic nouns with the structure CCVCC cannot be feminine or neuter, e.g. <i>der Stern</i>	Köpcke (1982: 84), Mills (1986: 33), Chan (2005: 296)	<i>Schwestertracht</i>
CCVV _{DIPHTHONG} - MONO→*F	Monosyllables beginning CCVV (where VV represents a diphthong) cannot be feminine, e.g. <i>das Blei</i>	Köpcke (1982: 86)	<i>Benediktenkraut, Dweil</i>
-CHEN→*M,*F	Nouns with the diminutive suffix <i>-chen</i> cannot be masculine or feminine, e.g. <i>das Bäumchen</i>	Köpcke (1982: 71), Mills (1986: 30), Steinmetz & Rice (1989: 165), Flämig (1991: 453), Hoberg (2004: 87), Menzel (2004: 66), Thomoglou (2004: 41), Weinrich (2007: 326), Duden Grammatik (2009: 166), Engel (2009: 281), Eisenberg (2013: 133)	<i>Goldhähnchen</i>

CONVERSION→*M,*F	Nominalised parts of speech cannot be masculine or feminine. This includes infinitives (e.g. <i>das Laufen</i>), adjectives (e.g. <i>das Schwarz</i>), adverbs (e.g. <i>das Hier und Jetzt</i>), pronouns (e.g. <i>das Ich</i>), conjunctions (e.g. <i>das Wenn und Aber</i>), prepositions (e.g. <i>das Auf und Ab</i>), interjections (e.g. <i>das Hallo</i>), and whole verb phrases (e.g. <i>das Inkrafttreten</i>)	Köpcke (1982: 74), Mills (1986: 30), Steinmetz (1986: 198), Flämig (1991: 453), Hickey (2000: 630), Hoberg (2004: 90f), Menzel (2004: 63), Chan (2005: 55ff.), Duden Grammatik (2009: 159), Engel (2009: 282), Eisenberg (2013: 134)	<i>Schaffen, Haben, Leben</i>
CV MONO→*N	Monosyllabic nouns with the structure CV cannot be neuter, e.g. <i>der Schuh</i>	Köpcke (1982: 83)	<i>Tourensch</i>
CVC MONO→*F	Monosyllabic nouns with the structure CVC cannot be feminine, e.g. <i>der Ton</i>	Köpcke (1982: 83)	<i>Sog, Dom</i>
CVCC MONO→*N	Monosyllabic nouns with the structure CVCC cannot be neuter, e.g. <i>der Schutz</i>	Köpcke (1982: 83)	<i>All-Star-Band, Windhund</i>
CVCC MONO→*F	Monosyllabic nouns with the structure CVCC cannot be feminine, e.g. <i>das Land</i>	Köpcke (1982: 83)	
CVCCC MONO→*F,*N	Monosyllabic nouns with the structure CVCCC cannot be feminine or neuter, e.g. <i>der Scherz</i>	Köpcke (1982: 83)	<i>Kunst, DAAD</i>
D- MONO→*F	Monosyllabic nouns whose onset is /d/ cannot be feminine, e.g. <i>der Dolch</i>	Köpcke (1982: 90)	<i>Dime, Dom</i>
DEVERBAL STEM→*F,*N	Nouns derived from bare verb stems cannot be feminine or neuter, e.g. <i>der Versuch</i> . This also includes ablauted verb stems, e.g. <i>der Gang</i> .	Flämig (1991: 453), Hickey (2000: 643), Hoberg (2004: 90), Chan (2005: 55)	<i>Zusammenbruch</i>
DEVERBAL -T→*M,*N	Nouns which have been derived from a verb stem + [t] cannot be masculine or neuter, e.g. <i>die Schrift</i>	Flämig (1991: 453), Hickey (2000: 634), Hoberg (2004: 87), Chan (2005: 55)	<i>Durchschrift</i>
DR-→*F,*N	Monosyllabic nouns beginning with the consonant cluster /dr/ cannot be feminine or neuter, e.g. <i>der Dreh</i>	Köpcke (1982: 88), Mills (1986: 33), Köpcke & Zubin (1997: 29), Menzel (2004: 68), Chan (2005: 296)	<i>Drall</i>
-E→*M,*N	Nouns ending in a schwa cannot be masculine or neuter, e.g. <i>die Fliege</i>	Altmann & Raettig (1973: 302), Mills (1986: 33), Steinmetz (1986: 192), Steinmetz & Rice (1989: 166), Wegener (1995), Köpcke & Zubin (1996: 476), Nelson (1998: 218), Hoberg (2004: 87), Rice (2006: 1396), Eisenberg (2013: 134)	<i>Hubbrücke, Palpe, Leiste</i>
-E.SUFFIX→*M,*N	Noun with the suffix -e cannot be masculine or neuter, e.g. <i>die Größe</i>	Mills (1986: 33), Steinmetz (1986: 192), Steinmetz & Rice (1989: 166), Wegener (1995), Köpcke & Zubin (1996: 476), Nelson (1998: 218), Hoberg (2004: 87), Eisenberg (2013: 134)	<i>Abwärme, Lese, Palpe</i>
-EDER→*M,*F	Nouns with the “foreign” suffix -eder cannot be masculine or feminine, e.g. <i>das Hexaeder</i>	Nelson (1998: 218)	<i>Hexaeder</i>
-EE/-ÉE/-É→*M,*F	Nouns with the “foreign” suffix -ee (incl. -ée and -é) cannot be masculine or feminine, e.g. <i>das Soufflé</i>	Fleischer & Barz (1992: 192), Bopp (2000-2018)	<i>Renforcé</i>

-EE/-ÉE/-É→*M,*N	Nouns with the “foreign” suffix <i>-ee</i> (incl. <i>-ée</i> and <i>-é</i>) cannot be masculine or neuter, e.g. <i>die Armee</i>	Fleischer & Barz (1992: 192), Bopp (2000-2018)	
-EE/-ÉE/-É→*F,*N	Nouns with the “foreign” suffix <i>-ee</i> (incl. <i>-ée</i> and <i>-é</i>) cannot be feminine or neuter, e.g. <i>der Frisée</i>	Bopp (2000-2018)	
-EHR/-ÄHR→*M,*N	Polysyllabic nouns ending in <i>-ehr</i> or <i>-ähr</i> cannot be masculine or neuter, e.g. <i>die Gewähr</i>	Steinmetz (2006: 1424)	<i>Containerverkehr</i>
-EKT→*F,*N	Nouns ending in <i>-ekt</i> [ekt] cannot be feminine or neuter, e.g. <i>der Architekt</i>	Fehring (2014: 2)	<i>Präfekt</i>
-EKT→*M,*F	Nouns ending in <i>-ekt</i> [ekt] cannot be masculine or feminine, e.g. <i>das Insekt</i>	Fehring (2014: 2)	
-EKT→*F	Nouns ending in <i>-ekt</i> [ekt] cannot be feminine, e.g. <i>der Intellekt, das Projekt</i>	Fehring (2014: 2)	
-EL→*F,*N	Nouns with the pseudo-suffix <i>-el</i> or suffix <i>-el</i> cannot be feminine or neuter, e.g. <i>der Flügel</i>	Steinmetz (1986: 197ff.), Flämig (1991: 453), Wegener (1995), Hoberg (2004: 92), Menzel (2004: 66), Durrell (2011: 9), Eisenberg (2013: 135)	<i>F-Schlüssel, Geisel, Nagel</i>
-EL→*M,*F	Nouns with the pseudo-suffix <i>-el</i> or suffix <i>-el</i> cannot be masculine or feminine, e.g. <i>das Segel</i>	Durrell (2011: 9)	
-EL→*M,*N	Nouns with the pseudo-suffix <i>-el</i> or suffix <i>-el</i> cannot be masculine or neuter, e.g. <i>die Trommel</i>	Steinmetz & Rice (1989: 166)	
-EN→*F,*N	Nouns with the pseudo-suffix <i>-en</i> cannot be feminine or neuter, e.g. <i>der Morgen</i>	Wegener (1995), Köpcke & Zubin (1996: 476), Hoberg (2004: 92), Menzel (2004: 66), Durrell (2011: 9), Eisenberg (2013: 135)	<i>Schaffen, Haben, Leben</i>
-EN→*M,*F	Nouns with the pseudo-suffix <i>-en</i> masculine or feminine, e.g. <i>das Zeichen</i>	Steinmetz & Rice (1989: 166), Hoberg (2004: 92)	
-EN→*F	Nouns with the pseudo-suffix <i>-en</i> cannot be feminine, e.g. <i>der Süden</i>	Steinmetz & Rice (1989: 166), Wegener (1995), Köpcke & Zubin (1996: 476), Menzel (2004: 66), Durrell (2011: 9), Eisenberg (2013: 135)	
-ENT→*F,*N	Nouns with the “foreign” suffix <i>-ent</i> cannot be feminine or neuter, e.g. <i>der Abiturient</i>	Flämig (1991: 453), Hoberg (2004: 86), Thomoglou (2004: 41), Weinrich (2007: 326), Hoberg (2004: 92), Engel (2009: 280)	<i>Suizident, Petent</i>
-ENT→*M,*F	Nouns ending in <i>-ent</i> cannot be masculine or feminine, e.g. <i>das Äquivalent</i>	Durrell (2011: 10)	
-ENZ→*M,*N	Nouns with the “foreign” suffix <i>-enz</i> cannot be masculine or neuter, e.g. <i>die Existenz</i>	Altmann & Raettig (1973: 302), Hoberg (2004: 87), Menzel (2004: 66), Weinrich (2007: 327) Duden Grammatik (2009: 164), Engel (2009: 280)	<i>Konvergenz, Renitenz</i>
-ER→*F,*N	Nouns with the (unstressed) pseudo-suffix <i>-er</i> cannot be feminine or neuter, e.g. <i>der Fehler</i>	Steinmetz (1986: 197), Wegener (1995), Hoberg (2004: 92), Durrell (2011: 9), Eisenberg (2013: 135)	<i>C-Dur-Tonleiter, Chargennummer</i>
-ER→*M,*F	Nouns with the (unstressed) pseudo-suffix <i>-er</i> cannot be masculine or feminine, e.g. <i>das Lager</i>	Durrell (2011: 9)	

-ER→*M,*N	Nouns with the (unstressed) pseudo-suffix <i>-er</i> cannot be masculine or neuter, e.g. <i>die Butter</i>	Steinmetz & Rice (1989: 166)	
-ER SUFF.→*F,*N	Nouns with the agentive/instrumental suffix <i>-er</i> cannot be feminine or neuter, e.g. <i>der Lehrer</i>	Flämig (1991: 453), Zifonun et al. (1997: 33), Hickey (2000: 635), Hoberg (2004: 86), Menzel (2004: 66), Thomoglou (2004: 41), Weinrich (2007: 326), Engel (2009: 280), Eisenberg (2013: 133)	<i>Topverdiener, Bergführer</i>
-ETT→*M,*F	Nouns ending in <i>-ett</i> cannot be masculine or feminine, e.g. <i>das Tablett</i>	Mills (1986: 33), Steinmetz & Rice (1989: 166), Köpcke & Zubin (1997), Hoberg (2004: 96), Menzel (2004: 68), Duden Grammatik (2009: 166)	<i>Triolett</i>
-ETTE→*M,*N	Nouns with the “foreign” suffix <i>-ette</i> cannot be masculine or neuter, e.g. <i>die Zigarette</i>	Thomoglou (2004: 41), Weinrich (2007: 327), Duden Grammatik (2009: 165), Engel (2009: 280)	<i>Toilette, Lafette</i>
-EUR/ÖR→*F,*N	Nouns with the “foreign” suffix <i>-eur/-ör</i> (incl. <i>-ateur, -iteur</i>) cannot be feminine or neuter, e.g. <i>der Friseur</i>	Flämig (1991: 453), Hoberg (2004: 86), Thomoglou (2004: 41), Weinrich (2007: 326), Engel (2009: 279)	<i>Galvaniseur</i>
-EUR/ÖR→*N	Nouns with the “foreign” suffix <i>-eur</i> cannot be neuter, e.g. <i>der Redakteur, die Couleur</i>	Engel (2009: 279)	
-FT/-CHT→*M,*N	Monosyllabic nouns ending in a non-sibilant fricative [f]/[x]/[ç] followed by [t] cannot be masculine or neuter, e.g. <i>die Frucht</i>	Zubin & Köpcke (1981: 440, 1984: 44), Köpcke (1982: 98), Mills (1986: 33), Köpcke & Zubin (1997), Hoberg (2004: 98), Menzel (2004: 68), Duden Grammatik (2009: 166), Eisenberg (2013: 135)	<i>Durchschrift</i>
GE→*M,*F	Nouns with the prefix <i>Ge-</i> cannot be masculine or feminine, e.g. <i>das Gefühl</i>	Mills (1986: 30), Hickey (2000: 631), Menzel (2004: 66), Rice (2006: 1396), Steinmetz (2006: 1424)	<i>Geleucht, Getändel</i>
-HEIT/-(IG)KEIT→*M,*N	Nouns with the suffix <i>-heit</i> or <i>-(ig)keit</i> cannot be masculine or neuter, e.g. <i>die Heiterkeit</i>	Zubin & Köpcke (1984: 44), Mills (1986: 30), Steinmetz & Rice (1989: 165), Flämig (1991: 453), Köpcke & Zubin (1996: 476), Zifonun et al. (1997: 32), Hickey (2000: 630), Hoberg (2004: 87), Menzel (2004: 66), Thomoglou (2004: 41), Weinrich (2007: 327), Duden Grammatik (2009: 165), Engel (2009: 281), Eisenberg (2013: 133)	<i>Einigkeit, Sachlichkeit</i>
-IE [i:]→*M,*N	Nouns with the stressed “foreign” suffix <i>-ie</i> [i:] cannot be masculine or neuter, e.g. <i>die Philosophie</i>	Altmann & Raettig (1973: 302), Flämig (1991: 453), Hoberg (2004: 87), Thomoglou (2004: 41), Weinrich (2007: 327), Duden Grammatik (2009: 165), Engel (2009: 281)	<i>Neotenie, Akrozephalie</i>
-IE [iə]→*M,*N	Nouns <i>-ie</i> [iə] cannot be masculine or neuter, e.g. <i>die Studie</i>	Altmann & Raettig (1973: 302), Hoberg (2004: 92), Thomoglou (2004: 41), Weinrich (2007: 327), Duden Grammatik (2009: 165)	<i>Ziellinie, Dahlie</i>
-IER /i:r/ POLY.→*M,*F	Polysyllabic nouns ending in stressed <i>-ier</i> [i:r] cannot be masculine or feminine, e.g. <i>das Klavier</i>	Mills (1986: 33), Steinmetz (1986: 197), Nelson (1998: 207, 212), Menzel (2004: 68)	<i>Rektapapier</i>
-IG→*F,*N	Polysyllabic nouns ending in <i>-ig</i> cannot be feminine or neuter, e.g. <i>der Honig</i>	Hoberg (2004: 96), Thomoglou (2004: 41), Duden Grammatik (2009: 164), Eisenberg (2013: 134)	<i>Beutenhonig</i>

-IG/-ICH→*F,*N	Nouns ending in [ɪç] (-ig/-ich) cannot be feminine or neuter, e.g. <i>der Käfig</i>	Altmann & Raettig (1973: 302), Mills (1986: 33), Hoberg (2004: 96), Duden Grammatik (2009: 164), Eisenberg (2013: 134)	<i>Beutenhonig</i>
-IK→*M,*N	Nouns with the “foreign” suffix -ik cannot be masculine or neuter, e.g. <i>die Mathematik</i>	Altmann & Raettig (1973: 302), Zifonun et al. (1997: 33), Nelson (1998: 218), Hoberg (2004: 87), Menzel (2004: 66), Thomoglou (2004: 41), Weinrich (2007: 327), Duden Grammatik (2009: 165), Engel (2009: 281)	<i>Aszetik, Thematik, Modallogik</i>
-ILLE→*M,*N	Nouns in -ille cannot be masculine or neuter, e.g. <i>die Kamille</i>	Duden Grammatik (2009: 165)	<i>Pupille</i>
-IN [i:n]→*M,*F	Nouns with the “foreign” suffix -in [i:n] cannot be masculine or feminine, e.g. <i>das Benzin</i>	Flämig (1991: 453), Duden Grammatik (2009: 166)	<i>Gerichtsmedizin</i>
-IN SUFF.→*M,*N	Nouns with the suffix -in cannot be masculine or neuter, e.g. <i>die Lehrerin</i>	Mills (1986: 30), Flämig (1991: 453), Hoberg (2004: 87), Thomoglou (2004: 41) Engel (2009: 281), Eisenberg (2013: 133)	<i>Generalagentin, Szientistin</i>
-INE→*M,*N	Nouns ending in -ine [i:nə] cannot be masculine or neuter, e.g. <i>die Maschine</i>	Weinrich (2007: 327), Duden Grammatik (2009: 165)	<i>Turbine</i>
-INER→*F,*N	Nouns with the suffix -iner cannot be feminine or neuter, e.g. <i>der Dalmatiner</i>	Bopp (2000-2018)	<i>Byzantiner</i>
-ING→*M,*F	Nouns (English loans) ending in the “foreign” suffix -ing cannot be masculine or feminine, e.g. <i>das Meeting</i>	Hoberg (2004: 87), Menzel (2004: 66), Duden Grammatik (2009: 166)	<i>Coaching</i>
-ION→*M,*N	Nouns with the “foreign” (stressed) suffix -ion (incl. -ation, -ition etc.) cannot be masculine or neuter, e.g. <i>die Dekoration</i>	Altmann & Raettig (1973: 302), Mills (1986: 30), Flämig (1991: 453), Nelson (1998: 217), Hickey (2000: 630), Hoberg (2004: 87), Menzel (2004: 66), Weinrich (2007: 327), Duden Grammatik (2009: 165), Engel (2009: 281)	<i>Disproportion, Kolmation</i>
-IS→*M,*N	Nouns ending in (unstressed) -is [ɪs] cannot be masculine or neuter, e.g. <i>die Praxis</i>	Nelson (1998: 216), Menzel (2004: 66), Duden Grammatik (2009: 165)	<i>Hybris</i>
-ISMUS→*F,*N	Nouns with the “foreign” suffix -ismus cannot be feminine or neuter, e.g. <i>der Kommunismus</i>	Altmann & Raettig (1973: 302), Zifonun et al. (1997: 33), Hoberg (2004: 86), Weinrich (2007: 326), Duden Grammatik (2009: 164), Engel (2009: 280)	<i>Konformismus</i>
-IST→*F,*N	Nouns with the “foreign” suffix -ist cannot be feminine or neuter, e.g. <i>der Sozialist</i>	Flämig (1991: 453), Hoberg (2004: 86), Thomoglou (2004: 41), Weinrich (2007: 326), Engel (2009: 280)	<i>Statist, Folklorist</i>
-IT→*F,*N	Nouns with the “foreign” suffix -it cannot be feminine or neuter, e.g. <i>der Israelit</i>	Bopp (2000-2018)	<i>Bronzit, Satellit, Metropolit</i>
-IT→*M,*F	Nouns with the “foreign” suffix -it cannot be masculine or feminine, e.g. <i>das Dynamit</i>	Bopp (2000-2018), Durrell (2011: 8)	
-IT→*F	Nouns with the “foreign” suffix -it cannot be feminine, e.g. <i>der Islamit</i>	Bopp (2000-2018), Durrell (2011: 8)	
-ITIS→*M,*N	Nouns with the “foreign” suffix -itis cannot be masculine or neuter, e.g. <i>die Dermatitis</i>	Weinrich (2007: 327), Duden Grammatik (2009: 165), Engel (2009: 281)	<i>Laryngitis</i>

-IUM→*M,*F	Nouns with the “foreign” suffix <i>-ium</i> (incl. <i>-arium</i>) cannot be masculine or feminine, e.g. <i>das Bakterium</i>	Altmann & Raettig (1973: 302), Flämig (1991: 453)	<i>Rubidium, Osmium, Vadium</i>
-IV→*N	Nouns ending in <i>-iv</i> [i:f] cannot be neuter, e.g. <i>der Nominativ</i>	Fehring (2014: 2)	<i>Motiv</i>
-IV→*M,*F	Nouns ending in <i>-iv</i> [i:f] cannot be masculine or feminine, e.g. <i>das Derivat</i>	Durrell (2011: 10)	
-IV→*F	Nouns ending in <i>-iv</i> [i:f] cannot be feminine, e.g. <i>das Adjektiv</i>	Nelson (1998: 217), Durrell (2011: 10), Fehring (2014: 2)	
-IV STRESSED→*M,*F	Nouns ending in stressed <i>-iv</i> [i:f] cannot be masculine or feminine, e.g. <i>das Adjektiv</i>	Nelson (1998: 217)	<i>Motiv</i>
-IVE→*M,*N	Nouns in <i>-ive</i> cannot be masculine or neuter, e.g. <i>die Perspektive</i>	Weinrich (2007: 327), Duden Grammatik (2009: 165)	<i>Exzellenzinitiative</i>
KR- MONO→*F,*N	Monosyllabic nouns with the onset /kr/ cannot be feminine or neuter, e.g. <i>der Kran</i>	Köpcke (1982: 91)	<i>Benediktenkraut</i>
-L MONO→*F	Monosyllabic nouns whose coda is /l/ cannot be feminine, e.g. <i>das Moll</i>	Köpcke (1982: 100)	<i>Drall, Reinfall</i>
-LING→*F,*N	Nouns with the suffix <i>-ling</i> cannot be feminine or neuter, e.g. <i>der Bratling</i>	Flämig (1991: 453), Köpcke & Zubin (1996: 476), Zifonun et al. (1997: 33), Hoberg (2004: 86), Menzel (2004: 66), Thomoglou (2004: 41), Weinrich (2007: 326), Duden Grammatik (2009: 164), Engel (2009: 280), Eisenberg (2013: 133)	<i>Pfifferling, Sechsling</i>
-LT MONO→*M,*F	Monosyllabic nouns ending in [lt] <i>-lt/-ld</i> cannot be masculine or feminine, e.g. <i>das Zelt</i>	Hickey (2000: 634)	<i>Kilovolt, Rauschgold</i>
-MA→*M,*F	Nouns (of Greek origin) with the “foreign” suffix <i>-ma</i> cannot be masculine or feminine, e.g. <i>das Thema</i>	Flämig (1991: 453), Weinrich (2007: 326), Duden Grammatik (2009: 166)	<i>Sophisma</i>
-MENT→*M,*F	Nouns with the “foreign” suffix <i>-ment</i> (incl. <i>-ament, -ement</i> and <i>-iment</i>) cannot be masculine or feminine, e.g. <i>das Testament</i>	Mills (1986: 30), Kratochvilová (2000: 71), Menzel (2004: 66), Steinmetz (1986: 197), Weinrich (2007: 326), Duden Grammatik (2009: 166), Engel (2009: 281)	<i>Indossament, Liniment</i>
MONOSYLL.→*F,*N	Monosyllabic nouns cannot be feminine or neuter, e.g. <i>der Traum</i>	Arndt (1970: 253), Carstensen (1980: 21f.), Köpcke (1982: 45), Wegener (1995: 87)	<i>Sog, Dom, Drall, Tran</i>
-N(C)(C) MONO→*F,*N	Monosyllabic nouns whose coda consists of a nasal (followed by up to two consonants) cannot be feminine or neuter, e.g. <i>der Grund</i>	Köpcke (1982: 100), Mills (1986: 33), Köpcke & Zubin (1997), Hoberg (2004: 98), Menzel (2004: 68)	<i>Schund, Windhund</i>
-NIS→*M,*F	Nouns with the suffix <i>-nis</i> cannot be masculine or feminine, e.g. <i>das Ereignis</i>	Steinmetz (1986: 200, 2006: 1424), Zifonun et al. (1997: 33) Hoberg (2004: 89), Thomoglou (2004: 41), Duden Grammatik (2009: 166), Engel (2009: 279), Eisenberg (2013: 133)	<i>Detailkenntnis, Finsternis</i>

-NIS→*M,*N	Nouns with the suffix <i>-nis</i> cannot be masculine or feminine, e.g. <i>die Erlaubnis</i>	Steinmetz (1986: 200, 2006: 1424), Zifonun et al. (1997: 33), Hoberg (2004: 89), Duden Grammatik (2009: 166), Engel (2009: 279)	
-NIS→*M	Nouns with the suffix <i>-nis</i> cannot be masculine, e.g. <i>die Besorgnis</i>	Steinmetz (1986: 200, 2006: 1424), Zifonun et al. (1997: 33), Hoberg (2004: 89), Thomoglou (2004: 41), Duden Grammatik (2009: 166), Engel (2009: 279), Eisenberg (2013: 133)	
-O→*M,*F	Nouns ending in <i>-o</i> [o:] (unstressed) cannot be masculine or feminine, e.g. <i>das Kino</i>	Hickey (1999: 641), Menzel (2004: 68)	<i>Allegro, Tempo</i>
-ON→*M,*F	Nouns ending in <i>-on</i> cannot be masculine or feminine, e.g. <i>das Lexicon</i>	Durrell (2011: 10), Bjornstad (2014: 78)	<i>Geison</i>
-ON UNSTRESS.→*M,*F	Nouns ending in the unstressed “foreign” suffix <i>-on</i> cannot be masculine or feminine, e.g. <i>das Stadion</i>	Nelson (1998: 217)	<i>Geison</i>
-OR→*F,*N	Nouns with the “foreign” suffix <i>-or</i> (incl. <i>-ator, -itor</i>) cannot be feminine or neuter, e.g. <i>der Detektor</i>	Flämig (1991: 453), Hoberg (2004: 86), Menzel (2004: 66), Thomoglou (2004: 41), Weinrich (2007: 326), Duden Grammatik (2009: 164), Engel (2009: 280)	<i>Thyristor</i>
∅V _{LONG} MONO→*F	Monosyllabic nouns with no onset and a long vowel cannot be feminine, e.g. <i>das Ohr</i>	Köpcke (1982: 86)	<i>Ar</i>
R- MONO→*F	Monosyllabic nouns whose onset is /r/ cannot be feminine, e.g. <i>der Rock</i>	Köpcke (1982: 90)	<i>Rähm</i>
-RIS→*M,*N	Polysyllabic nouns ending in unstressed <i>-ris</i> cannot be masculine or neuter, e.g. <i>die Iris</i>	Nelson (1998: 216)	<i>Hybris</i>
-SAL→*M	Nouns with the suffix <i>-sal</i> cannot be masculine, e.g. <i>das Rinnsal</i>	Flämig (1991: 453), Duden Grammatik (2009: 166)	<i>Mühsal</i>
-SAL→*M,*F	Nouns with the suffix <i>-sal</i> cannot be masculine or feminine, e.g. <i>das Scheusal</i>	Flämig (1991: 453), Duden Grammatik (2009: 166)	
-SAL→*M,*N	Nouns with the suffix <i>-sal</i> cannot be masculine or neuter, e.g. <i>die Trübsal</i>	Flämig (1991: 453), Duden Grammatik (2009: 166)	
-SCHAFT→*M,*N	Nouns with the suffix <i>-schaft</i> cannot be masculine or neuter, e.g. <i>die Freundschaft</i>	Mills (1986: 30), Flämig (1991: 453), Zifonun et al. (1997: 33), Hickey (2000: 630), Hoberg (2004: 87), Menzel (2004: 66), Thomoglou (2004: 41), Weinrich (2007: 327), Duden Grammatik (2009: 165), Engel (2009: 281), Eisenberg (2013: 133)	<i>Rechtswissenschaft</i>
-SKOP→*M,*F	Nouns with the “foreign” suffix <i>-skop</i> cannot be masculine or feminine, e.g. <i>das Periskop</i>	Flämig (1991: 453)	<i>Polarisationsmikroskop</i>
-ST/-SCHT/-ZT→*N	Monosyllabic nouns ending in a sibilant fricative [s]/[ʃ] followed by [t] cannot be neuter, e.g. <i>die Pest</i>	Köpcke (1982: 97)	<i>Kost, Kirschegeist, Test</i>
-ST/-SCHT/-ZT→*F,*N	Monosyllabic nouns ending in a sibilant fricative [s]/[ʃ] followed by [t] cannot be feminine or neuter, e.g. <i>der Trost</i>	Köpcke (1982: 97)	

T- MONO→*F	Monosyllabic nouns whose onset is /t/ cannot be feminine, e.g. <i>der Tod</i>	Köpcke (1982: 90)	<i>Test</i>
-TÄT→*M,*N	Nouns with the “foreign” suffix <i>-(i)tät</i> cannot be masculine or neuter, e.g. <i>die Universität</i>	Mills (1986: 30), Flämig (1991: 453), Zifonun et al. (1997: 33), Hoberg (2004: 87), Menzel (2004: 66), Thomoglou (2004: 41), Weinrich (2007: 327), Duden Grammatik (2009: 165), Engel (2009: 281)	<i>Elektrizität, Pauschalität</i>
-THEK→*M,*N	Nouns in <i>-(o)thek</i> cannot be masculine or neuter, e.g. <i>Diskothek</i>	Nelson (1998: 219)	<i>Österreichische Nationalbibliothek</i>
TR→*F,*N	Monosyllabic nouns beginning with the consonant cluster /tr/ cannot be feminine or neuter, e.g. <i>der Traum</i>	Köpcke (1982: 88), Mills (1986: 33), Köpcke & Zubin (1997: 29), Menzel (2004: 68), Chan (2005: 296)	<i>Tran</i>
-TUM→*F,*N	Nouns with the suffix <i>-tum</i> cannot be feminine or neuter, e.g. <i>der Reichtum</i>	Menzel (2004: 66)	<i>Anachoretentum</i>
-TUM→*M,*F	Nouns with the suffix <i>-tum</i> cannot be masculine or feminine, e.g. <i>das Heiligtum</i>	Flämig (1991: 453), Nelson (1998: 215), Hickey (2000: 629), Hoberg (2004: 87), Thomoglou (2004: 41), Duden Grammatik (2009: 166), Engel (2009: 282), Eisenberg (2013: 133)	
-TUM→*F	Nouns with the suffix <i>-tum</i> cannot be feminine, e.g. <i>das Christentum</i>	Flämig (1991: 453), Nelson (1998: 215), Hickey (2000: 629), Hoberg (2004: 87), Menzel (2004: 66), Thomoglou (2004: 41), Duden Grammatik (2009: 166), Engel (2009: 282), Eisenberg (2013: 133)	
-UM→*M,*F	Nouns with the unstressed “foreign” suffix <i>-um</i> or <i>-ikum</i> cannot be masculine or feminine, e.g. <i>das Praktikum</i>	Steinmetz (1985, 1986: 198), Flämig (1991: 453), Nelson (1998: 215), Menzel (2004: 66), Weinrich (2007: 326), Duden Grammatik (2009: 166)	<i>Supremum, Ferrum</i>
-UNG→*M,*N	Nouns with the suffix <i>-ung</i> cannot be masculine or neuter, e.g. <i>die Meinung</i>	Altmann & Raettig (1973: 302), Mills (1986: 30), Flämig (1991: 453), Köpcke & Zubin (1996: 476), Zifonun et al. (1997: 33), Hickey (2000: 630), Hoberg (2004: 87), Menzel (2004: 66), Thomoglou (2004: 41), Weinrich (2007: 327), Duden Grammatik (2009: 165), Engel (2009: 281), Eisenberg (2013: 133)	<i>Bloßstellung, Verzögerung</i>
-UR SUFF.→*M,*N	Nouns with the “foreign” suffix <i>-ur</i> (incl. <i>-atur</i> and <i>-itur</i>) cannot be masculine or neuter, e.g. <i>die Architektur</i>	Flämig (1991: 453), Köpcke & Zubin (1997), Menzel (2004: 66), Hoberg (2004: 87), Thomoglou (2004: 41), Weinrich (2007: 327), Duden Grammatik (2009: 165), Engel (2009: 280)	<i>Rezeptur</i>
-UR STRESSED→*M,*N	Nouns ending in stressed /'u:r/ (<i>-ur/-uhr</i>) cannot be masculine or neuter, e.g. <i>die Nomenklatur</i>	Mills (1986: 33), Hickey (2000: 631), Menzel (2004: 68)	<i>Rezeptur</i>
-ÜR/-ÜHR /y:r/→*M,*N	Nouns ending in /y:r/ (<i>-ür/-ühr</i>) cannot be masculine or neuter, e.g. <i>die Gebühr</i>	Mills (1986: 33), Hickey (2000: 631), Menzel (2004: 68), Chan (2005: 296)	<i>Porphy</i>
-US→*F,*N	Nouns ending in <i>-us</i> cannot be feminine or neuter, e.g. <i>der Rhythmus</i>	Menzel (2004: 66), Weinrich (2007: 326)	<i>Konformismus</i>

VC MONO→*M,*F	Monosyllabic nouns with the structure VC cannot be masculine or feminine, e.g. <i>das Ohr</i>	Köpcke (1982: 83)	<i>Ar</i>
VCC MONO→*F,*N	Monosyllabic nouns with the structure VCC cannot be feminine or neuter, e.g. <i>der Ast</i>	Köpcke (1982: 83)	<i>Postgiroamt</i>
VCCC MONO→*M	Monosyllabic nouns with the structure VCCC cannot be masculine, e.g. <i>die Angst</i>	Köpcke (1982: 83)	<i>Examensangst</i>
VER-→*F	Nouns with the prefix <i>Ver-</i> cannot be feminine, e.g. <i>das Verbot</i>	Steinmetz (2006: 1424)	<i>Verbesserer, Verleitung</i>
VER- NO SUFF.→*F	Nouns with the prefix <i>Ver-</i> which do not have a suffix cannot be feminine, e.g. <i>der Versuch</i>	Steinmetz (2006: 1424)	<i>Verdruss, Verkauf</i>
V _{LONG} MONO→*F	Monosyllabic nouns containing a long vowel cannot be feminine, e.g. <i>der Sohn</i>	Zubin & Köpcke (1981: 440), Köpcke (1982: 95)	<i>Dom, Sog, Tran</i>
V _{LONG} C MONO→*F	Monosyllabic nouns containing a long vowel and a consonant in the coda cannot be feminine, e.g. <i>der Mohn</i>	Köpcke (1982: 87)	<i>Dom, Sog, Tran</i>
V _{SHORT} STOP C→*F,*N	Monosyllabic nouns containing a short vowel followed by a stop and another consonant cannot be feminine or neuter, e.g. <i>der Platz</i>	Köpcke (1982: 103)	<i>Hops, Interrogativsatz</i>
V _{STRESSED} [i]→*M,*F	Polysyllabic nouns ending in a stressed vowel followed by [i] cannot be masculine or feminine, e.g. <i>das Profil</i>	Nelson (1998: 214)	<i>Epistyl</i>
V _{STRESSED} [m]→*M,*F	Polysyllabic nouns ending in a stressed vowel followed by [m] cannot be masculine or feminine, e.g. <i>das Problem</i>	Nelson (1998: 215)	<i>Erschöpfungssyndrom</i>
VV _{DIPHTH} C MONO→*F	Monosyllabic nouns containing a diphthong and a consonant in the coda cannot be feminine, e.g. <i>der Laut</i>	Köpcke (1982: 88)	<i>Dweil, Langhaus</i>
ACAD. DISCIPLINES→*M,*N	Nouns denoting academic disciplines cannot be masculine or neuter, e.g. <i>die Germanistik, die Kunst</i>	Flämig (1991: 452), Köpcke & Zubin (1996: 480), Hickey (2000: 629), Menzel (2004: 63), Weinrich (2007: 327), Engel (2009: 272)	<i>Kunst</i>
ALC. DRINKS→*F,*N	Nouns denoting alcoholic drinks cannot be feminine or neuter, e.g. <i>der Wein, der Schnaps</i>	Zubin & Köpcke (1981: 444), Köpcke (1982: 72), Zubin & Köpcke (1984: 44), Köpcke & Zubin (1996: 479), Hickey (2000: 630), Menzel (2004: 63), Thomoglou (2004: 34), Chan (2005: 94), Duden Grammatik (2009: 159), Engel (2009: 272), Eisenberg (2013: 138)	<i>Brandy, Kirschegeist</i>
ANG. SHAPES→*M,*F	Nouns denoting shapes with angular geometric features cannot be masculine or feminine, e.g. <i>das Parallelogram</i>	Nelson (1998: 218)	<i>Hexaeder</i>
ANNOUNCE→*M,*F	Nouns denoting announcements/ pronouncements of an official or public nature cannot be masculine or feminine, e.g. <i>das Bulletin, das Verdikt</i>	Nelson (1998: 219)	<i>Äußerung, Erkundigung</i>
BIRDS→*F,*N	Nouns denoting birds cannot be feminine or neuter, e.g. <i>der Adler</i>	Mills (1986: 28), Köpcke & Zubin (1996)	<i>Goldhähnchen, Stelzvögel</i>

BOATS→*F,*N	Nouns denoting hand or motor craft cannot be feminine or neuter, e.g. <i>der Kajak</i>	Mills (1986: 28)	<i>Achter</i>
BODIES OF WATER→*N	Nouns denoting expanses of water cannot be neuter, e.g. <i>der Teich, der Strom, die Lagune</i>	Köpcke (1982: 77)	<i>Lorenzstrom</i>
CHEESES→*F,*N	Nouns denoting types of cheese cannot be feminine or neuter, e.g. <i>der Brie, der Feta, der Gorgonzola</i>	Hoberg (2004: 109)	<i>Holländer</i>
CHEMICAL COMP./SUBST.→*M,*F	Nouns referring to chemical compounds/substances cannot be masculine or feminine, e.g. <i>das Insulin, das Salz</i>	Flämig (1991: 452), Nelson (1998: 213)	<i>Zyan</i>
CITIES/TOWNS→*M,*F	Names of cities or towns cannot be masculine or feminine, e.g. <i>das schöne Paris</i>	Hickey (2000: 630), Hoberg (2004: 106), Menzel (2004: 63), Thomoglou (2004: 34), Chan (2005: 96), Köpcke & Zubin (2005a: 119), Weinrich (2007: 328), Engel (2009: 272), Duden Grammatik (2009: 159), Eisenberg (2013: 138)	<i>Antwerpen, Münster</i>
CONDUCT→*M,*N	Nouns denoting abstract principles as guides to conduct cannot be masculine or neuter, e.g. <i>die Doktrin, die Regel</i>	Nelson (1998: 218)	<i>Aszetik, Regel</i>
COUNTRIES→*M,*F	Names of countries cannot be masculine or feminine, e.g. <i>das Deutschland, das China</i>	Hickey (2000: 630), Hoberg (2004: 107), Thomoglou (2004: 34), Chan (2005: 96), Weinrich (2007: 328), Duden Grammatik (2009: 159), Engel (2009: 272)	<i>Großbritannien</i>
DAYS→*F,*N	Names of days cannot be feminine or neuter, e.g. <i>der Montag, der Freitag</i>	Zubin & Köpcke (1984: 44), Flämig (1991: 452), Köpcke & Zubin (1996: 480), Hickey (2000: 629), Menzel (2004: 63), Chan (2005: 94), Weinrich (2007: 327), Engel (2009: 272), Eisenberg (2013: 138)	<i>Martinstag</i>
DISEASES→*M,*N	Nouns denoting diseases or disorders cannot be masculine or neuter, e.g. <i>die Polio, die Impetigo</i>	Nelson (1998: 218)	<i>Akrozephalie, Lyssa, Laryngitis</i>
DOM. ANIMALS→*M,*F	Nouns denoting domesticated animals cannot be masculine or feminine, e.g. <i>das Schwein, das Schaf</i>	Zubin & Köpcke (1981: 444), Köpcke (1982: 76), Mills (1986: 27), Hoberg (2004: 102), Duden Grammatik (2009: 157), Eisenberg (2013: 138)	<i>Sikahirsch, Windhund</i>
DWELL→*M,*F	Nouns denoting dwellings cannot be masculine or feminine, e.g. <i>das Haus, das Hotel, das Chalet</i>	Onysko (2007: 173)	<i>Substandardwohnung</i>
ELAB. INSTR.→*M,*N	Nouns denoting instrumental objects with a functionally elaborated end, e.g. furcated or pincer-like objects, cannot be masculine or neuter, e.g. <i>die Gabel, die Zange</i>	Mills (1986: 27), Steinmetz (1986: 205ff.), Steinmetz & Rice (1989: 166)	<i>Nuteisen</i>
ELEMENTS→*M,*F	Nouns denoting chemical elements cannot be masculine or feminine, e.g. <i>das Argon, das Zink</i>	Köpcke (1982: 74), Mills (1986: 27), Flämig (1991: 452), Köpcke & Zubin (1996: 480), Menzel (2004: 63), Thomoglou (2004: 34), Weinrich (2007: 328)	<i>Osmium, Ferrum</i>
EXTROVERSION→*F,*N	Nouns denoting extroversion cannot be feminine or neuter, e.g. <i>der Zorn, der Mut, der Eifer</i>	Zubin & Köpcke (1984: 44), Mills (1986: 28), Köpcke & Zubin (1997)	<i>Hybris</i>
F. HOLLOWES→*M,*F	Nouns denoting functional hollows or openings cannot be masculine or feminine, e.g. <i>das Auge, das Rad, das Tor</i>	Steinmetz (2006: 1434)	<i>Aufzugsschacht, Orgelpfeife</i>

FABRIC→*F,*N	Nouns denoting types of fabric or cloth cannot be feminine or neuter, e.g. <i>der Samt</i>	Mills (1986: 27), Menzel (2004: 63)	<i>Knopflochseide, Juchtenleder</i>
FEAR→*M,*N	Nouns denoting fear, anxiety or urgency cannot be masculine or neuter, e.g. <i>die Angst, die Not</i>	Steinmetz (1985: 20)	<i>Kabinettskrise, Übereile</i>
FISH→*F,*N	Nouns denoting fish are masculine, e.g. <i>der Dorsch</i>	Köpcke & Zubin (1996)	<i>Rapfen</i>
FLOWERS→*M,*N	Names of flowers cannot be masculine or neuter, e.g. <i>die Rose, die Nelke, die Tulpe</i>	Flämig (1991: 452), Hickey (2000: 630), Menzel (2004: 63), Thomoglou (2004: 34), Chan (2005: 95), Weinrich (2007: 328)	<i>Bärentraube, Dahlie</i>
FRUITS→*M,*N	Nouns denoting fruits (and nuts) cannot be masculine or neuter, e.g. <i>die Kiwi, die Mango, die Birne</i>	Köpcke & Zubin (1996: 480), Duden Grammatik (2009: 158)	<i>Cranberry, Bärentraube</i>
GAMES→*M,*F	Nouns denoting games or sports cannot be masculine or feminine, e.g. <i>das Hockey, das Schach, das Badminton</i>	Mills (1986: 27), Köpcke & Zubin (1996: 480), Chan (2005: 96)	<i>Blackjack, Seniorengymnastik</i>
GEN. PERSON→*F,*N	Nouns denoting a generic person (without specifying their societal gender) cannot be feminine or neuter, e.g. <i>der Mensch, der Gast, der Monarch</i>	Köpcke (1982: 72), Mills (1986: 27), Hoberg (2004: 103), Chan (2005: 91ff.), Onysko (2007: 159)	<i>Bautzener, Bergführer, Präfekt, Presenter</i>
GEN. PERSON→*M,*F	Nouns referring to unspecified individuals cannot be masculine or feminine, e.g. <i>das Individuum, das Mitglied</i>	Flämig (1991: 456)	<i>Bautzener, Bergführer, Präfekt, Presenter</i>
GESTURES→*M,*N	Nouns denoting gestures cannot be masculine or neuter, e.g. <i>die Gebärde</i>	Mills (1986: 27)	<i>Gebärdenspiel</i>
GROUPS→*M,*N	Nouns denoting collectives of individuals cannot be masculine or neuter, e.g. <i>die Gruppe, die Mannschaft</i>	Onysko (2007: 164)	<i>Komplexbrigade, Expertengruppe</i>
GROUPS→*M,*F	Nouns denoting collectives of individuals cannot be masculine or feminine, e.g. <i>das Volk, das Publikum</i>	Flämig (1991: 456)	<i>Komplexbrigade, Expertengruppe</i>
HUNTING→*M,*N	Nouns belonging to the semantic field of hunting cannot be masculine or neuter, e.g. <i>die Jagd</i>	Mills (1986: 27)	<i>Grandel, Kelle</i>
IMPROPER SUPERORDINATES→*M,*F	Nouns denoting 'improper superordinates' cannot be masculine or feminine. Improper superordinates 'represent a purely mental classification' requiring 'judgements of a human mind', e.g. <i>das Ziel, das Rätsel, das Wunder</i>	Steinmetz (1986: 201, 2006: 1433), Nelson (1998: 220)	<i>(Detailkenntnis, Thematik)</i>
INSECTS→*M,*N	Nouns denoting insects cannot be masculine or neuter, e.g. <i>die Fliege, die Spinne</i>	Mills (1986: 28), Köpcke & Zubin (1996)	<i>Federling</i>
INSTITUTIONS→*M,*F	Nouns denoting institutions cannot be masculine or feminine, e.g. <i>das Internat, das Krankenhaus</i>	Onysko (2007: 173)	<i>Krankenanstalt</i>
INTEGRATED PARTS→*F,*N	Nouns denoting integrated parts cannot be feminine or neuter, e.g. <i>der Bestandteil</i>	Mills (1986: 28)	<i>Läufer, Rähm</i>
INTROVERSION→*M,*N	Nouns denoting introversion cannot be masculine or neuter, e.g. <i>die Scheu, die Geduld, die Anmut</i>	Zubin & Köpcke (1984), Mills (1986: 28), Köpcke & Zubin (1997), Eisenberg (2013: 134)	<i>Konivenz, Wortkargheit</i>
ISLANDS→*M,*F	Names of islands cannot be masculine or feminine, e.g. <i>das Malta, das Rügen</i>	Chan (2005: 96), Engel (2009: 272)	<i>Großbritannien</i>

LANGUAGES→*M,*F	Names of languages cannot be masculine or feminine, e.g. <i>das Deutsch, das Schwedisch</i>	Köpcke (1982: 74), Flämig (1991: 453), Köpcke & Zubin (1996: 480), Menzel (2004: 63) Chan (2005: 96)	<i>Farsi</i>
LEATHER TYPES→*M,*F	Nouns denoting types of leather cannot be masculine or feminine, e.g. <i>das Chamois</i>	Schulte-Beckhausen (2002: 73), Chan (2005: 47)	<i>Juchten, Juchtenleder</i>
LETTERS→*M,*F	Nouns denoting letters of the alphabet cannot be masculine or feminine, e.g. <i>das A, das B</i>	Köpcke (1982: 74), Flämig (1991: 451), Hoberg (2004: 90), Menzel (2004: 64), Thomoglou (2004: 34), Chan (2005: 96), Weinrich (2007: 329), Duden Grammatik (2009: 159), Engel (2009: 272), Durrell (2011: 2)	<i>J, Q</i>
LIVING/WORKING→*M,*F	Nouns denoting living or working spaces cannot be masculine or feminine, e.g. <i>das Haus, das Quartier</i>	Nelson (1998: 207), Kraiss (2014: 36)	<i>Bootcamp, Kantonbank, Langhaus</i>
MALE ANIMALS→*F,*N	Nouns denoting male animals cannot be feminine or neuter, e.g. <i>der Bulle, der Hengst, der Hahn</i>	Köpcke (1982: 76), Mills (1986: 27), Flämig (1991: 451), Köpcke & Zubin (1996: 484), Hoberg (2004: 100), Weinrich (2007: 336), Duden Grammatik (2009: 157), Eisenberg (2013: 137), Bjornstad (2014: 76)	<i>Farre</i>
MAMMALS→*F,*N	Nouns denoting mammals cannot be feminine or neuter, e.g. <i>der Hund, der Elefant</i>	Köpcke & Zubin (1996)	<i>Sikahirsch, Windhund, Mustang</i>
MEN→*F,*N	Nouns denoting men cannot be feminine or neuter, e.g. <i>der Mann, der König, der Onkel</i>	Zubin & Köpcke (1981: 445), Köpcke (1982: 75), Mills (1986: 27), Flämig (1991: 451), Köpcke & Zubin (1996: 479), Hoberg (2004: 99), Weinrich (2007: 331), Duden Grammatik (2009: 154), Eisenberg (2013: 137)	<i>Eunuch, Knacker</i>
MESH→*M,*F	Nouns denoting mesh-like structures or cruciform objects cannot be masculine or feminine, e.g. <i>das Netz, das Sieb</i>	Nelson (1998: 219), Onysko (2007: 173)	<i>Rotang, Garnknäuel</i>
METALS→*M,*F	Nouns denoting metals cannot be masculine or feminine, e.g. <i>das Gold, das Eisen, das Zinn</i>	Köpcke (1982: 74), Menzel (2004: 63), Thomoglou (2004: 34), Weinrich (2007: 328), Eisenberg (2013: 138)	<i>Osmium, Ferrum, Rubidium, Nuteisen</i>
MON. UNITS→*F,*N	Nouns denoting monetary units cannot be feminine or neuter, e.g. <i>der Euro, der Dollar, der Forint</i>	Menzel (2004: 63), Thomoglou (2004: 34), Chan (2005: 94)	<i>Rupiah, Dime</i>
MOUNTAINS→*F,*N	Names of mountains cannot be feminine or neuter, e.g. <i>der Ben Nevis, der Kilamandscharo</i>	Flämig (1991: 452), Chan (2005: 94), Weinrich (2007: 328), Duden Grammatik (2009: 160f.)	<i>Nanga Parbat</i>
MUS. COMP.→*M,*F	Nouns denoting types of musical compositions cannot be masculine or feminine, e.g. <i>das Allegro, das Konzert</i>	Schulte-Beckhausen (2002: 72)	<i>Allegro</i>
MUS. INSTR.→*M,*F	Nouns denoting musical instruments cannot be masculine or feminine, e.g. <i>das Horn, das Saxophon</i>	Steinmetz (1986: 197), Nelson (1998: 219)	<i>Tamtam, Bassgeige</i>
MUS. INSTR.→*M,*N	Nouns denoting musical instruments cannot be masculine or neuter, e.g. <i>die Oboe, die Gitarre</i>	Mills (1986: 27)	<i>Tamtam, Bassgeige</i>
MUS. INSTR.→*M	Nouns denoting musical instruments cannot be masculine, e.g. <i>die Orgel, die Sitar, das Klavier</i>	Mills (1986: 27), Steinmetz (1986: 197), Nelson (1998: 219)	<i>Tamtam, Bassgeige</i>
MUS. NOTES→*M,*F	Nouns denoting musical notes cannot be masculine or feminine, e.g. <i>das C, das Cis, das Ces</i>	Köpcke (1982: 75), Nelson (1998: 219), Engel (2009: 272), Durrell (2011: 2)	<i>His</i>
MUS. TERMS→*M,*F	Nouns denoting musical terms cannot be masculine or feminine, e.g. <i>das Crescendo, das Dur, das Oktav</i>	Nelson (1998: 219)	<i>Allegro, C-Dur-Tonleiter, His</i>

MUSIC TYPES→*F,*N	Nouns denoting types of music cannot be feminine or neuter, e.g. <i>der Pop, Jazz, Hip-Hop</i>	Onysko (2007: 159)	<i>Janitscharenmusik, New-Orleans-Jazz</i>
NAME: MAN→*F,*N	Names of men cannot be feminine or neuter, e.g. <i>der Johannes, der Karl, der Müller</i>	Duden Grammatik (2009: 154)	<i>Walter, Tintoretto, Lärtes, Schiele, Iskariot, Klimt, Naftali</i>
NAME: WOMAN→*M,*N	Names of women cannot be masculine or neuter, e.g. <i>die Steffi, die Eva, die Klum</i>	Duden Grammatik (2009: 154)	<i>Nadja, Sophia, Rahel, Sigrun</i>
NAUTICAL→*M,*N	Nouns belonging to the semantic field of nautical language cannot be masculine or neuter, e.g. <i>die Crew, die Drift</i>	Köpcke (1982: 78)	<i>Flettnerruder, Kimm</i>
NO INFO.→*M,*F	Nouns which provide no specific information about their referent cannot be masculine or feminine, e.g. <i>das Ding, das Objekt, das Zeug</i>	Zubin & Köpcke (1986: 144), Köpcke & Zubin (1997)	<i>Fremdstoff</i>
OPEN WATER→*M,*N	Nouns denoting open bodies of water cannot be masculine or neuter, e.g. <i>die See</i>	Mills (1986: 28)	<i>Lorenzstrom</i>
OPENINGS→*M,*F	Nouns denoting openings cannot be masculine or feminine. This constraint is related to FUNCTIONAL HOLLOWES. 'In addition to open spaces serving as passages or gaps, or holes and apertures, 'openings' here denotes also cavities or depressions', e.g. <i>das Becken, Fenster, Grab</i>	Nelson (1998: 219)	<i>Rinne</i>
ORGANS→*M,*F	Nouns denoting major organs cannot be masculine or feminine, e.g. <i>das Herz, das Auge</i>	Steinmetz (1986: 209)	<i>Magenschleimhaut</i>
PCC GARMENT→*M,*F	Nouns denoting primary chest-covering garments cannot be masculine or feminine. The garment must cover the area between the waist and the chest and must not be outerwear. E.g. <i>das Kleid, das Hemd</i>	Steinmetz (1985: 20, 1986: 195), Steinmetz & Rice (1989: 166)	<i>Zweiteiler</i>
PICTURES→*M,*F	Nouns denoting pictorial images cannot be masculine or feminine, e.g. <i>das Bild, das Foto, das Image</i>	Onysko (2007: 173)	<i>Diaphanbild</i>
PLANTS→*F,*N	Nouns denoting plants cannot be feminine or neuter, e.g. <i>der Kaktus</i>	Mills (1986: 28)	<i>Benediktenkraut, Pimpernuss</i>
POWER/STRENGTH→*M,*N	Nouns denoting power and strength cannot be masculine or neuter, e.g. <i>die Gewalt</i>	Mills (1986: 27)	<i>Vertretungsvollmacht</i>
PREP. DISHES→*M,*F	Nouns denoting prepared dishes cannot be masculine or feminine, e.g. <i>das Chop-suey, das Ragout, das Soufflé</i>	Nelson (1998: 218f.)	<i>Eclair, Hasenpfeffer</i>
PRIMATES→*F,*N	Nouns denoting primates are masculine, e.g. <i>der Gorilla</i>	Köpcke & Zubin (1996)	<i>Kapuzineraffe</i>
REGIONS→*M,*F	Nouns denoting regions, provinces, states etc. cannot be masculine or feminine, e.g. <i>das Bayern</i>	Hickey (2000: 630), Hoberg (2004: 106), Menzel (2004: 63), Thomoglou (2004: 34), Köpcke & Zubin (2005a: 119), Weinrich (2007: 328) Duden Grammatik (2009: 159), Engel (2009: 272), Eisenberg (2013: 138)	<i>Patagonien, Bessarabien</i>
RIVERS EUROPE→*M,*N	Names of European rivers cannot be masculine or neuter, e.g. <i>die Seine</i>	Nübling (2015: 316)	<i>Garonne</i>

RIVERS NON-EUROPE→*F,*N	Names of non-european rivers cannot be feminine or neuter, e.g. <i>der Amazonas</i>	Nübling (2015: 316)	<i>Lorenzstrom</i>
RIVERS NON-GER.→*F,*N	Nouns denoting non-German rivers cannot be feminine or neuter, e.g. <i>der Ganges, der Jordan, der Nil</i>	Flämig (1991: 451), Duden Grammatik (2009: 161)	<i>Garonne, Lorenzstrom</i>
ROCKS/MINERALS →*F,*N	Nouns denoting rocks and minerals cannot be feminine or neuter, e.g. <i>der Diamant, der Quarz</i>	Köpcke (1982: 71f.), Mills (1986: 27), Flämig (1991: 452), Menzel (2004: 63), Thomoglou (2004: 34), Chan (2005: 94), Weinrich (2007: 328), Engel (2009: 272)	<i>Bronzit, Porphy</i>
SCI. UNITS→*M,*F	Nouns denoting scientific units and units of measurement cannot be masculine or feminine, e.g. <i>das Atom, das Proton, das Watt, das Kilo</i>	Köpcke (1982: 74), Eisenberg (1999: 155), Hoberg (2004: 106), Menzel (2004: 63), Thomoglou (2004: 34), Flämig (1991: 453), Chan (2005: 95)	<i>Fahrenheit, Kilovolt, Mikrofarad, Monomer</i>
SEASONS→*F,*N	Nouns denoting seasons cannot be feminine or neuter, e.g. <i>der Sommer, der Herbst</i>	Flämig (1991: 452), Köpcke & Zubin (1996: 480), Hickey (2000: 629), Menzel (2004: 63), Chan (2005: 94), Weinrich (2007: 327), Engel (2009: 272)	<i>Sommer</i>
SEMANTIC DIMINUTIVES→*M,*N	Nouns denoting 'semantic diminutives' cannot be masculine or neuter. Semantic diminutives are said to refer to 'smaller, weaker, less active and less prominent' entities, e.g. <i>die Insel (vs. das Land), die Tür (vs. das Tor), die Matte (vs. der Teppich)</i>	Steinmetz (1986: 202f.), Steinmetz & Rice (1989: 166)	<i>Rinne (vs. e.g. Schlucht)</i>
SEPARABLE PARTS→*M,*F	Nouns denoting separable parts cannot be masculine or feminine, e.g. <i>das Einzelteil</i>	Mills (1986: 28)	<i>Läufer, Rähm</i>
SHARP→*M,*N	Nouns denoting objects with a sharp form cannot be masculine or neuter, e.g. <i>die Gabel</i>	Mills (1986: 27)	<i>Nagel, Klinge</i>
SHEET/STRIP/SLAB→*M,*F	Nouns denoting sheets, strips or slabs cannot be masculine or feminine, e.g. <i>das Blatt, das Blech, das Brett</i>	Steinmetz (2001: 219f.)	<i>Rauschgold, Leiste</i>
SPEECH→*M,*N	Nouns denoting acts of speech cannot be masculine or neuter, e.g. <i>die Frage, die Rede</i>	Mills (1986: 27), Menzel (2004: 63)	<i>Antwort, Äußerung</i>
STICK→*F,*N	Nouns denoting stick-like objects with an instrumental function cannot be feminine or neuter, e.g. <i>der Dolch, der Stift</i>	Steinmetz (1983), Zubin & Köpcke (1984: 45), Mills (1986: 27)	<i>Löffelstiel, Nagel</i>
SUP.→*M,*F	Nouns denoting superordinate categories cannot be masculine or feminine, e.g. <i>das Obst, das Tier, das Spiel</i>	Mills (1986: 27), Steinmetz (1986: 192, 2006: 1424), Zubin & Köpcke (1986), Eisenberg (1999: 156), Rice (2006: 1398), Duden Grammatik (2009: 158)	<i>Himmelskörper, Kunst</i>
SUPERORDINATES OF INDIRECT REFERENCE→*M,*F	Nouns denoting 'superordinates of indirect reference' cannot be masculine or feminine. These are defined by 'vagueness of the denotata'. The category is said to 'reflect a basic property of neuter nouns in general (highlighted by Zubin & Köpcke 1984: 144) that neuter nouns index lesser differentiation in their referents and categorize objects at the highest level of generalisation', e.g. <i>Erlebnis, Risiko, Bedürfnis, Desaster, Detail, Scheusal</i>	Nelson (1998: 221)	<i>(Detailkenntnis, Geisel)</i>

TEMPERATURES→*M,*N	Nouns denoting temperatures cannot be masculine or neuter, e.g. <i>Hitze, Glut</i>	Mills (1986: 27), Menzel (2004: 63)	<i>Abwärme</i>
TIME: SHORT→*F,*N	Nouns denoting short periods of time cannot be feminine or neuter, e.g. <i>Tag</i>	Mills (1986: 27)	<i>Martinstag</i>
TIMES OF DAY→*F,*N	Nouns denoting times of day cannot be feminine or neuter, e.g. <i>der Morgen, der Abend</i>	Menzel (2004: 63), Thomoglou (2004: 34)	<i>Abend, Skatabend</i>
TREES→*M,*N	Nouns denoting trees cannot be masculine or neuter, e.g. <i>die Zeder, die Eiche, die Buche</i>	Flämig (1991: 452), Hickey (2000: 629), Menzel (2004: 63), Thomoglou (2004: 34), Chan (2005: 95), Weinrich (2007: 328), Engel (2009: 272)	<i>Affenbrotbaum, Mandelbaum</i>
UNITS OF TIME→*F,*N	Nouns denoting units of time cannot be feminine or neuter, e.g. <i>der Tag</i>	Köpcke (1982: 71)	<i>Karwoche, Martinstag</i>
VEG. ROOT→*M,*N	Nouns denoting vegetables which consist of the root of the plant cannot be masculine or neuter, e.g. <i>die Rübe</i>	Köpcke & Zubin (2005a: 152)	<i>Salatkartoffel</i>
WASTE→*F,*N	Nouns denoting rubbish and waste cannot be feminine or neuter, e.g. <i>Dreck, Müll</i>	Mills (1986: 27)	<i>Schund</i>
WOMEN→*M,*N	Nouns denoting women cannot be masculine or neuter, e.g. <i>die Frau, die Schwester, die Tante</i>	Zubin & Köpcke (1981: 445), Köpcke (1982: 75), Mills (1986: 27), Steinmetz & Rice (1989: 165), Flämig (1991: 451), Köpcke & Zubin (1996: 481), Hoberg (2004: 99), Weinrich (2007: 331), Duden Grammatik (2009: 154), Eisenberg (2013: 137)	<i>Holländerin, Szientistin, Enkelin</i>
WORD CLASSES→*M,*F	Nouns denoting word classes cannot be masculine or feminine, e.g. <i>Verb, Adjektiv</i>	Köpcke (1982: 73), Nelson (1998: 218), Menzel (2004: 63), Chan (2005: 95)	<i>Nomen proprium</i>
YOUNG ANIMALS→*M,*F	Nouns denoting young animals cannot be masculine or feminine, e.g. <i>das Ferkel, Fohlen, Kalb, Lamm</i>	Zubin & Köpcke (1981: 444) Köpcke (1982: 77), Flämig (1991: 456), Eisenberg (1999: 156), Hickey (2000: 630), Hoberg (2004: 102), Chan (2005: 296), Weinrich (2007: 336), Duden Grammatik (2009: 157)	<i>Farre</i>
ACCESS PASSAGES→*M,*N	Nouns denoting access passages cannot be masculine or neuter, e.g. <i>die Gangway</i>	Steinmetz & Rice (1989), Onysko (2007: 177)	---
APPLES→*F,*N	Types of apple cannot be feminine or neuter, e.g. <i>der Cox</i>	Hoberg (2004: 109)	---
BEER→*M,*F	Nouns denoting types of beer cannot be masculine or feminine, e.g. <i>das Pilsner</i>	Mills (1986: 27), Hoberg (2004: 108), Duden Grammatik (2009: 159)	---
CAFÉS→*M,*F	Names of cafés cannot be masculine or feminine, e.g. <i>das Kranzler, das Kaufmann</i>	Hoberg (2004: 108), Menzel (2004: 64), Thomoglou (2004: 34), Chan (2005: 94), Weinrich (2007: 329), Duden Grammatik (2009: 162), Engel (2009: 272)	---
CARBONATED DRINKS→*M,*N	Nouns denoting carbonated drinks cannot be masculine or neuter, e.g. <i>die Brause</i>	Mills (1986: 27)	---
CARS→*F,*N	Nouns denoting makes of car cannot be feminine or neuter, e.g. <i>der Audi, der BMW</i>	Flämig (1991: 452), Hickey (2000: 630), Hoberg (2004: 108), Menzel (2004: 63), Köpcke & Zubin (2005a),	---

		Weinrich (2007: 328), Duden Grammatik (2009: 162), Engel (2009: 272)	
CIGARETTES→*M,*N	Nouns denoting brands of cigarettes and cigars cannot be masculine or neuter, e.g. <i>die Camel, die Lord</i>	Hoberg (2004: 108), Menzel (2004: 63), Chan (2005: 94), Engel (2009: 272)	---
CINEMAS→*M,*F	Names of cinemas cannot be masculine or feminine, e.g. <i>das Roxy, das Corona</i>	Hickey (2000: 630), Menzel (2004: 64), Hoberg (2004: 108), Thomoglou (2004: 34), Chan (2005: 94), Köpcke & Zubin (2005a: 114), Weinrich (2007: 329) Duden Grammatik (2009: 162)	---
COINS/NOTES→*F,*N	Nouns denoting denominations of coins and banknotes derived from numbers cannot be feminine or neuter, e.g. <i>der Fünfer, Fünziger</i>	Engel (2009: 272)	---
COLOURS→*M,*F	Nouns denoting colours cannot be masculine or feminine, e.g. <i>das Blau, das Schwarz</i>	Mills (1986: 27), Köpcke & Zubin (1996: 480), Köpcke & Zubin (1997: 45f.), Chan (2005: 94), Weinrich (2007: 328), Engel (2009: 272)	---
COMPANIES→*F	Names of companies cannot be feminine, e.g. ‚Nestlé hat sein...‘; ‚Apple hat sein...‘	Duden Grammatik (2009: 162)	---
COMPASS POINTS→*F,*N	Nouns denoting points of the compass cannot be feminine or neuter, e.g. <i>der Norden, der Osten</i>	Köpcke (1982: 71), Flämig (1991: 452), Menzel (2004: 63), Thomoglou (2004: 34), Chan (2005: 94), Weinrich (2007: 327), Engel (2009: 272), Eisenberg (2013: 138)	---
CONTINENTS→*M,*F	Names of continents cannot be masculine or feminine, e.g. <i>das Afrika, das Europa</i>	Hoberg (2004: 107), Menzel (2004: 63), Thomoglou (2004: 34), Chan (2005: 94), Weinrich (2007: 328), Engel (2009: 272)	---
CUSHIONS/RECLINING→*M,*F	Nouns related to cushions or reclining cannot be masculine or feminine, e.g. <i>das Bett, das Kissen</i>	Steinmetz (2001), Onysko (2007: 170)	---
DETERGENT→*M,*F	Names of detergents cannot be masculine or feminine, e.g. <i>das Persil, das Spüli</i>	Hoberg (2004: 108), Menzel (2004: 64), Chan (2005: 94), Engel (2009: 272)	---
DIMENSIONS→*M,*N	Nouns denoting dimensions cannot be masculine or neuter, e.g. <i>die Größe, die Breite</i>	Fehringer (2014: 3)	---
DRUGS→*M,*F	Nouns denoting (illegal) drugs cannot be masculine or feminine, e.g. <i>das Kokain, das Ecstasy</i>	Onysko (2007: 173)	---
EMANATIONS→*F,*N	Nouns denoting emanations cannot be feminine or neuter. An emanation has been defined by Steinmetz (1986: 201) as ‘something which originates from a source’, e.g. <i>der Geruch, der Gedanke</i>	Steinmetz (1986: 201)	---
ENCLOSED WATER→*F,*N	Nouns denoting enclosed bodies of water cannot be feminine or neuter, e.g. <i>der See</i>	Mills (1986: 28)	---

FEMALE ANIMALS→*M,*N	Nouns denoting female animals cannot be masculine or neuter, e.g. <i>die Gans, die Henne, die Eselin</i>	Köpcke (1982: 76), Mills (1986: 27), Flämig (1991: 451), Köpcke & Zubin (1996: 484), Hoberg (2004: 100), Weinrich (2007: 336), Duden Grammatik (2009: 157), Eisenberg (2013: 137)	---
GERMAN RIVERS→*M,*N	Nouns denoting German rivers cannot be masculine or neuter, e.g. <i>die Donau, die Elbe</i>	Flämig (1991: 452), Duden Grammatik (2009: 161)	---
GRASS→*M,*F	Nouns denoting types of grass cannot be masculine or feminine, e.g. <i>das Heu</i>	Mills (1986: 28)	---
HERBS/SPICES→*F,*N	Nouns denoting names of herbs and spices cannot be feminine or neuter, e.g. <i>der Kümmel</i>	Mills (1986: 28)	---
HOTELS→*M,*F	Names of hotels cannot be masculine or feminine, e.g. <i>das Hilton, das Adlon</i>	Hickey (2000: 630), Hoberg (2004: 108), Menzel (2004: 64), Thomoglou (2004: 34), Chan (2005: 94), Köpcke & Zubin (2005a: 114), Weinrich (2007: 329), Duden Grammatik (2009: 162)	---
MEDICATION→*M,*F	Nouns denoting medicines cannot be masculine or feminine, e.g. <i>das Aspirin</i>	Menzel (2004: 64), Thomoglou (2004: 34)	---
MEN: PEJ.→*M,*N	Nouns denoting pejorative terms for men 'who traditionally have been regarded in German culture as lacking central characteristics of culturally-defined masculinity' (Zubin & Köpcke 1981: 445), cannot be masculine or neuter, e.g. <i>die Tunte</i>	Köpcke & Zubin (1996: 483), Chan (2005: 296)	
MOLLUSCS→*M,*N	Nouns denoting molluscs cannot be masculine or neuter, e.g. <i>die Schnecke</i>	Mills (1986: 28), Köpcke & Zubin (1996)	---
MONTHS→*F,*N	Names of months cannot be feminine or neuter, e.g. <i>der Januar, der Februar</i>	Flämig (1991: 452), Köpcke & Zubin (1996: 480), Hickey (2000: 629), Menzel (2004: 63), Chan (2005: 94), Weinrich (2007: 327), Engel (2009: 272)	---
MOTORBIKES→*M,*N	Nouns denoting makes of motorbike cannot be masculine or neuter, e.g. <i>die BMW, die Honda</i>	Flämig (1991: 452), Hoberg (2004: 108), Menzel (2004: 63), Weinrich (2007: 328) Duden Grammatik (2009: 162), Engel (2009: 272), Eisenberg (2013: 138)	---
MOUNTAIN RANGES→*F,*N	Names of mountain ranges cannot be feminine or neuter, e.g. <i>der Harz, der Himalaja</i>	Flämig (1991: 452), Chan (2005: 94), Duden Grammatik (2009: 161)	---
NON-ALCOHOLIC DRINKS→*F,*N	Nouns denoting non-alcoholic drinks cannot be feminine or neuter, e.g. <i>der Saft, der Kaffee</i>	Duden Grammatik (2009: 159)	---
NUMBERS→*M,*N	Nouns denoting cardinal numbers cannot be masculine or neuter, e.g. <i>die Eins, die Zwei</i>	Köpcke (1982: 73), Mills (1986: 27), Flämig (1991: 451), Menzel (2004: 63), Thomoglou (2004: 34), Weinrich (2007: 328) Engel (2009: 272)	---

PETROL→*M,*F	Types of petrol cannot be masculine or feminine, e.g. <i>das Diesel, das Super</i>	Hoberg (2004: 109)	---
PHONEMES→*M,*F	Phonemes cannot be masculine or feminine, e.g. <i>das /t/</i>	Hoberg (2004: 90)	---
PLANES→*M,*N	Names of aeroplanes cannot be masculine or neuter, e.g. <i>die Boeing, die Spitfire</i>	Flämig (1991: 451), Hickey (2000: 630), Chan (2005: 94), Weinrich (2007: 328), Duden Grammatik (2009: 161f.), Engel (2009: 272)	---
PLANETS→*F,*N	Nouns denoting planets and other celestial bodies cannot be feminine or neuter, e.g. <i>der Mars, der Saturn</i>	Mills (1986: 27), Menzel (2004: 63)	---
PLANT-BASED DRINKS→*F,*N	Nouns denoting plant-based drinks cannot be feminine or neuter, e.g. <i>der Kako, der Tee</i>	Durrell (2011: 2)	---
PRECIPITATION→*F,*N	Nouns denoting forms of precipitation and types of weather cannot be feminine or neuter, e.g. <i>der Regen, der Schnee</i>	Köpcke (1982: 71), Mills (1986: 27), Flämig (1991: 452), Menzel (2004: 63), Thomoglou (2004: 34), Chan (2005: 94), Weinrich (2007: 327) Engel (2009: 272)	---
PREDATORS→*F,*N	Nouns denoting predators cannot be feminine or neuter, e.g. <i>der Tiger</i>	Schwichtenberg & Schiller (2004: 330)	---
PRODUCTION SITES→*M,*N	Nouns denoting production and processing sites cannot be masculine or neuter, e.g. <i>die Fabrik</i>	Steinmetz (1985: 20)	---
RESTAURANTS→*M,*F	Names of restaurants cannot be masculine or feminine, e.g. <i>das Siena</i>	Menzel (2004: 64), Köpcke & Zubin (2005a: 114), Duden Grammatik (2009: 162)	---
RIVERS IN SW GERMANY→*F,*N	Rivers in southwest Germany cannot be feminine or neuter, e.g. <i>der Neckar</i>	Bjornstad (2014: 77)	---
SEAFARING SIGNS→*M,*N	Nouns denoting seafaring signs cannot be masculine or neuter, e.g. <i>die Boje</i>	Mills (1986: 27)	---
SHAPE: SURFACE→*M,*N	Nouns denoting objects with a large, flat surface cannot be masculine or neuter, e.g. <i>die Tafel</i>	Mills (1986: 27)	---
SHIPS→*M,*N	Names of ships cannot be masculine or neuter, e.g. <i>die Bismarck, die Cutty Sark</i>	Mills (1986: 28), Flämig (1991: 451), Hickey (2000: 630), Hoberg (2004: 108), Menzel (2004: 63), Thomoglou (2004: 34), Chan (2005: 94), Köpcke & Zubin (2005a: 119), Weinrich (2007: 328), Duden Grammatik (2009: 161), Engel (2009: 272), Eisenberg (2013: 138)	---
SHOES→*F,*N	Nouns denoting types of shoe cannot be feminine or neuter, e.g. <i>der Stiefel, der Stöckel</i>	Onysko (2007: 173)	---
SHORELINES→*M,*F	Nouns denoting shorelines cannot be masculine or feminine, e.g. <i>das Ufer, das Kap</i>	Steinmetz (1985: 20, 2001: 219)	---
SIMPLE BLADE→*M,*F	Nouns denoting instrumental objects with a simple blade cannot be masculine or feminine, e.g. <i>das Messer, das Beil</i>	Steinmetz (1986: 205), Steinmetz & Rice (1989: 166)	---

SNAKES→*M,*N	Nouns denoting snakes cannot be masculine or neuter, e.g. <i>die Anakonda</i>	Mills (1986: 28), Köpcke & Zubin (1996)	---
SPICES→*F,*N	Nouns denoting spices cannot be feminine or neuter, e.g. <i>der Zimt</i>	Schwichtenberg & Schiller (2004: 330)	
TEA/COFFEE→*F,*N	Nouns denoting types of tea and coffee cannot be feminine or neuter, e.g. <i>der Darjeeling, der Arabica</i>	Menzel (2004: 63), Chan (2005: 94), Engel (2009: 272)	---
TIDEWATER→*M,*F	Nouns denoting bodies of tidal water cannot be masculine or feminine, e.g. <i>das Fleet</i>	Mills (1986: 28)	---
TIME SPANS→*M,*N	Nouns denoting time spans cannot be masculine or neuter, e.g. <i>die Brut</i>	Mills (1986: 27)	---
TIME: LONG→*M,*F	Nouns denoting longer periods of time cannot be masculine or feminine, e.g. <i>das Jahrzehnt</i>	Mills (1986: 27)	---
TIME: SHORTEST→*M,*N	Nouns denoting very short periods of time cannot be masculine or neuter, e.g. <i>die Sekunde</i>	Mills (1986: 27)	---
TRAINS→*F,*N	Names of trains cannot be feminine or neuter, e.g. <i>der ICE, der Regionalexpress</i>	Hoberg (2004: 109), Menzel (2004: 63), Engel (2009: 272)	---
TROPICAL FRUIT→*M,*N	Nouns denoting tropical fruits cannot be masculine or neuter, e.g. <i>die Papaya</i>	Steinmetz (1986: 192)	---
VEG. LEAF→*F,*N	Nouns denoting vegetables which consist of the leaf of the plant cannot be feminine or neuter, e.g. <i>der Spinat</i>	Köpcke & Zubin (2005a: 152)	---
VEG. STEM→*F,*N	Nouns denoting vegetables which consist of the stem of the plant cannot be feminine or neuter, e.g. <i>der Rhabarber</i>	Köpcke & Zubin (2005a: 152)	---
VEG. SEED/POD→*M,*N	Nouns denoting vegetables which consist of the seeds or pods of the plant cannot be masculine or neuter, e.g. <i>die Erbse</i>	Köpcke & Zubin (2005a: 152)	---
WET AREAS→*M,*N	Nouns denoting wet areas cannot be masculine or neuter, e.g. <i>die Pfütze</i>	Mills (1986: 28)	---
WINDS→*F,*N	Nouns denoting types of winds cannot be feminine or neuter, e.g. <i>der Föhn, Passat, Hurrikan</i>	Köpcke (1982: 71), Flämig (1991: 452), Hoberg (2004: 109), Menzel (2004: 63), Thomoglou (2004: 34), Chan (2005: 94), Engel (2009: 272), Eisenberg (2013: 138)	---
WINE→*F,*N	Names of wines cannot be feminine or neuter, e.g. <i>der Chardonnay, Riesling, Rioja</i>	Hoberg (2004: 108), Duden Grammatik (2009: 159), Engel (2009: 272)	---
WOMEN: PEJ.→*M,*F	Nouns denoting pejorative terms for women cannot be masculine or feminine, e.g. <i>das Luder</i>	Köpcke & Zubin (1996: 483), Chan (2005: 296)	---
YOUNG HUMANS→*M,*F	Nouns referring to young humans cannot be masculine or feminine, e.g. <i>das Baby, das Kind</i>	Zubin & Köpcke (1981: 445), Flämig (1991: 456), Hickey (2000: 630), Chan (2005: 296)	---

-AGOGE→*F,*N	Nouns with the “foreign” suffix <i>-agoge</i> cannot be feminine or neuter, e.g. <i>der Demagoge, der Pädagoge</i>	Engel (2009: 279)	---
-AILLE→*M,*N	Nouns in <i>-aille</i> cannot be masculine or neuter, e.g. <i>die Journaille, die Kanaille</i>	Duden Grammatik (2009: 164)	---
-AISE/-ÄSE→*M,*N	Nouns in <i>-aise</i> or <i>-äse</i> cannot be masculine or neuter, e.g. <i>die Mayonnaise/Majonäse, Polonaise/Polonäse</i>	Duden Grammatik (2009: 164)	---
-ANCE→*M,*N	Nouns ending in <i>-ance</i> cannot be masculine or feminine, e.g. <i>die Renaissance</i>	Duden Grammatik (2009: 164)	---
-AND→*F,*N	Nouns with the “foreign” suffix <i>-and</i> cannot be feminine or neuter, e.g. <i>der Doktorand, der Proband</i>	Duden Grammatik (2009: 164)	---
-ÄNE→*M,*N	Nouns in <i>-äne</i> cannot be masculine or neuter, e.g. <i>die Quarantäne, die Fontäne, die Moräne</i>	Hoberg (2004: 86), Weinrich (2007: 326), Engel (2009: 279)	---
-ANE→*F,*N	Nouns with the “foreign” suffix <i>-ane</i> cannot be feminine or neuter, e.g. <i>der Birmane</i>	Duden Grammatik (2009: 164)	---
-ANS→*M,*F	Nouns with the “foreign” suffix <i>-ans</i> cannot be masculine or feminine, e.g. <i>das Laxans</i>	Bopp (2000-2018)	---
-ÄR→*M,*F	Nouns in <i>-är</i> (which do not refer to people) cannot be masculine or feminine, e.g. <i>das Militär</i>	Bopp (2000-2018)	---
-ÄR→*F,*N	Nouns with the “foreign” suffix <i>-är</i> cannot be feminine or neuter, e.g. <i>der Revolutionär</i>	Flämig (1991: 453), Hoberg (2004: 86), Weinrich (2007: 326), Engel (2009: 280)	---
-ASE→*M,*N	Nouns with the “foreign” suffix <i>-ase</i> cannot be masculine or feminine, e.g. <i>die Hydratase, die Laktase</i>	Duden Grammatik (2009: 166)	---
-ASMUS→*F,*N	Nouns with the “foreign” suffix <i>-asmus</i> cannot be feminine or neuter, e.g. <i>der Sarkasmus</i>	Altmann & Raettig (1973: 302), Flämig (1991: 453), Hoberg (2004: 86), Weinrich (2007: 326), Engel (2009: 280)	---
-AST→*F,*N	Nouns in <i>-ast</i> cannot be feminine or neuter, e.g. <i>der Enthusiast</i>	Hoberg (2004: 89), Engel (2009: 280)	---
-ÄT→*M,*N	Nouns in <i>-ät</i> cannot be masculine or neuter, e.g. <i>die Diät</i>	Hoberg (2004: 89), Durrell (2011: 10), Bjornstad (2014: 78)	---
-ATE→*M,*N	Nouns with the “foreign” suffix <i>-ate</i> cannot be masculine or feminine, e.g. <i>die Koordinate</i>	Nelson (1998: 216)	---
-AZEE→*M,*N	Nouns with the “foreign” suffix <i>-azee</i> cannot be masculine or neuter, e.g. <i>die Rosasazee, die Oleazee</i>	Bopp (2000-2018)	---
-ELLE→*M,*N	Nouns ending in <i>-elle</i> cannot be masculine or neuter, e.g. <i>die Frikadelle</i>	Duden Grammatik (2009: 164)	---
-EM→*F,*N	Nouns ending in <i>-em</i> cannot be feminine or neuter, e.g. <i>der Atem</i>	Eisenberg (2013: 135)	---

-END→*F,*N	Nouns with the “foreign” suffix <i>-end</i> cannot be feminine or neuter, e.g. <i>der Subtrahend</i>	Bopp (2000-2018)	---
-ENS→*F,*N	Nouns with the “foreign” suffix <i>-ens</i> cannot be feminine or neuter, e.g. <i>das Detergens</i>	Bopp (2000-2018)	---
-ERIE→*M,*N	Nouns with the “foreign” suffix <i>-erie</i> cannot be masculine or neuter, e.g. <i>die Drogerie, die Parfümerie</i>	Hoberg (2004: 87), Engel (2009: 280)	---
-ESE→*F,*N	Nouns with the “foreign” suffix <i>-ese</i> cannot be feminine or neuter, e.g. <i>der Sudanese</i>	Bopp (2000-2018)	---
-ESS(E)→*M,*N	Nouns with the “foreign” suffix <i>-ess(e)</i> cannot be masculine or neuter, e.g. <i>die Hostess(e), Stewardess</i>	Flämig (1991: 453), Weinrich (2007: 327), Engel (2009: 280)	---
-EUSE→*M,*N	Nouns with the “foreign” suffix <i>-euse</i> cannot be masculine or neuter, e.g. <i>die Friseur</i>	Weinrich (2007: 327), Engel (2009: 281)	---
-EX→*F,*N	Nouns ending in <i>-ex</i> cannot be feminine or neuter, e.g. <i>der Komplex</i>	Nelson (1998: 218)	---
-GON→*M,*F	Nouns ending in <i>-gon</i> cannot be masculine or feminine, e.g. <i>das Hexagon</i>	Nelson (1998: 218)	---
GR- MONO→*F,*N	Monosyllabic nouns with the onset /gr/ cannot be feminine or neuter, e.g. <i>der Griff</i>	Köpcke (1982: 91)	---
-GRAMM→*M,*F	Nouns in <i>-gramm</i> cannot be masculine or feminine, e.g. <i>das Telegramm</i>	Nelson (1998: 219)	---
-IA→*M,*N	Nouns with the “foreign” suffix <i>-ia</i> cannot be masculine or neuter, e.g. <i>die Dahlia</i>	Bopp (2000-2018)	---
-ICHT→*M,*F	Nouns ending in <i>-icht</i> cannot be masculine or feminine, e.g. <i>das Licht, das Gewicht</i>	Flämig (1991: 453), Engel (2009: 282)	---
-ID→*M,*F	Nouns with the “foreign” suffix <i>-id</i> cannot be masculine or feminine, e.g. <i>das Chlorid</i>	Bopp (2000-2018)	---
-IER [je:]→*F,*N	Nouns with the “foreign” suffix <i>-ier</i> [je:] cannot be feminine or neuter, e.g. <i>der Bankier, Hotelier</i>	Hoberg (2004: 86), Thomoglou (2004: 41), Weinrich (2007: 326)	---
-IER [i:r]→*F,*N	Nouns with the “foreign” suffix <i>-ier</i> [i:r] cannot be feminine or neuter, e.g. <i>der Kanonier, der Juwelier</i>	Bopp (2000-2018)	---
-IERE→*M,*N	Nouns with the “foreign” suffix <i>-iere</i> cannot be masculine or neuter, e.g. <i>die Garderobiere</i>	Bopp (2000-2018)	---
-IKER→*F,*N	Nouns with the “foreign” suffix <i>-iker</i> cannot be feminine or neuter, e.g. <i>der Agnostiker, der Alkoholiker</i>	Zifonun et al. (1997: 33), Weinrich (2007: 326), Engel (2009: 280)	---
-IL→*M,*F	Nouns in <i>-il</i> cannot be masculine or feminine, e.g. <i>das Ventil</i>	Steinmetz (1986: 197)	---
-ISSE→*M,*N	Nouns with the “foreign” suffix <i>-isse</i> cannot be masculine or neuter, e.g. <i>die Kulisse, Prämisse</i>	Weinrich (2007: 327), Duden Grammatik (2009: 165)	---

-ITER→*F,*N	Nouns with the “foreign” suffix <i>-iter</i> cannot be feminine or neuter, e.g. <i>der Samariter</i>	Bopp (2000-2018)	---
-IV UNSTRESSED→*F,*N	Nouns ending in unstressed <i>-iv</i> [i:f] cannot be feminine or neuter, e.g. <i>der Dativ</i>	Nelson (1998: 217)	---
KN- MONO→*F,*N	Monosyllabic nouns with the onset /kn/ cannot be feminine or neuter, e.g. <i>der Knast, der Knopf</i>	Köpcke (1982: 89), Mills (1986: 33), Köpcke & Zubin (1997), Hoberg (2004: 97), Menzel (2004: 68)	---
-LCH→*F,*N	Nouns ending in /lx/ cannot be feminine or neuter, e.g. <i>der Dolch</i>	Hickey (2000: 634)	---
-LEIN→*M,*F	Nouns with the suffix <i>-lein</i> cannot be masculine or feminine, e.g. <i>das Büchlein</i>	Köpcke (1982: 71), Mills (1986: 30), Steinmetz & Rice (1989: 165), Flämig (1991: 453), Hoberg (2004: 87), Menzel (2004: 66), Thomoglou (2004: 41), Weinrich (2007: 326) Duden Grammatik (2009: 166), Engel (2009: 281), Eisenberg (2013: 133)	---
-LER→*F,*N	Nouns with the suffix <i>-ler</i> cannot be feminine or neuter, e.g. <i>der Wissenschaftler</i>	Flämig (1991: 453), Zifonun et al. (1997: 33), Hoberg (2004: 86), Eisenberg (2013: 133)	---
-LOG→*F,*N	Nouns in <i>-log</i> cannot be feminine or neuter, e.g. <i>der Dialog</i>	Altmann & Raettig (1973: 302)	---
-NER→*F,*N	Nouns with the suffix <i>-ner</i> cannot be feminine or neuter, e.g. <i>der Rentner</i>	Flämig (1991: 453), Hoberg (2004: 86)	---
-NESS→*M,*N	Nouns with the English suffix <i>-ness</i> cannot be masculine or neuter, e.g. <i>die Fitness</i>	Duden Grammatik (2009: 165)	---
-NIS STATE OF MIND→*M,*N	Deverbal nouns with the suffix <i>-nis</i> denoting a state of mind cannot be masculine or neuter, e.g. <i>die Besorgnis</i>	Durrell (2011: 10)	---
-OID→*M,*F	Nouns in <i>-oid</i> cannot be masculine or feminine, e.g. <i>das Trapezoid</i>	Nelson (1998: 218)	---
-OIR→*M,*F	Nouns ending in (stressed) <i>-oir</i> cannot be masculine or feminine, e.g. <i>das Trottoir</i>	Nelson (1998: 216)	---
-OL→*M,*F	Nouns with the “foreign” suffix <i>-ol</i> cannot be masculine or feminine, e.g. <i>das Methanol, das Menthol</i>	Bopp (2000-2018)	---
-OLCH→*F,*N	Nouns ending in <i>-olch</i> cannot be feminine or neuter, e.g. <i>der Dolch, der Molch</i>	Hickey (2000: 634)	---
-OM→*M,*F	Nouns with the “foreign” suffix <i>-om</i> cannot be masculine or feminine, e.g. <i>das Leprom, Neurom</i>	Bopp (2000-2018)	---
-ON STRESSED→*F,*N	Nouns ending in stressed <i>-on</i> cannot be feminine or neuter, e.g. <i>der Balkon</i>	Nelson (1998: 217)	---
-OSE→*M,*N	Nouns with the “foreign” suffix <i>-ose</i> cannot be masculine or neuter, e.g. <i>die Tuberkulose, Osmose</i>	Weinrich (2007: 327), Duden Grammatik (2009: 165)	---

-OST→*F,*N	Monosyllabic nouns ending in <i>-ost</i> cannot be feminine or neuter, e.g. <i>der Most</i>	Hickey (2000: 643)	---
-PF MONO→*F,*N	Monosyllables ending in <i>-pf</i> cannot be feminine or neuter, e.g. <i>der Knopf</i>	Hickey (2000: 645)	---
-PHON→*M,*F	Nouns in <i>-(o)phon</i> cannot be masculine or feminine, e.g. <i>das Mikrophon</i>	Flämig (1991: 453)	---
-PT→*M,*F	Nouns ending in <i>-pt</i> cannot be masculine or feminine, e.g. <i>das Rezept</i>	Nelson (1998: 219)	---
-SEL→*M,*F	Nouns with the suffix <i>-sel</i> cannot be masculine or feminine, e.g. <i>das Anhängsel</i>	Flämig (1991: 453), Hoberg (2004: 87)	---
-SIS UNSTRESSED→*M,*N	Nouns ending in unstressed <i>-sis</i> cannot be masculine or neuter, e.g. <i>die Basis, die Dosis</i>	Onysko (2007: 164), Weinrich (2007: 327), Duden Grammatik (2009: 166)	---
-TEL→*M,*F	Nouns with the suffix <i>-tel</i> cannot be masculine or feminine, e.g. <i>das Viertel</i>	Thomoglou (2004: 41)	---
-UCHT→*M,*N	Nouns ending in <i>-ucht</i> cannot be masculine or neuter, e.g. <i>die Bucht</i>	Steinmetz (1986: 192), Nelson (1998: 218), Rice (2006: 1398)	---
-UHR→*M,*N	Nouns ending in <i>-uhr</i> cannot be masculine or neuter, e.g. <i>die Abfuhr</i>	Mills (1986: 33), Hickey (2000: 631), Menzel (2004: 68)	---
-UKT→*M,*F	Nouns ending in <i>-ukt</i> cannot be masculine or feminine, e.g. <i>das Viadukt</i>	Nelson (1998: 219)	---
-UNFT→*M,*N	Nouns ending in <i>-unft</i> cannot be masculine or neuter, e.g. <i>die Vernunft</i>	Hickey (2000: 631)	---
-ÜRE→*M,*N	Nouns ending in <i>-üre</i> cannot be masculine or neuter, e.g. <i>die Broschüre</i>	Duden Grammatik (2009: 165)	---
-UST→*M,*N	Nouns ending in <i>-ust</i> cannot be masculine or neuter, e.g. <i>die Lust</i>	Hickey (2000: 643)	---
-V _{LONG+HIGH} /R/ MONO→*M,*N	Monosyllables containing a long and high vowel (/i:/, /u:/ or /y:/) followed by /r/ cannot be masculine or neuter, e.g. <i>die Gier</i>	Köpcke (1982: 103)	---
V MONO→*M	Monosyllabic nouns with the structure V cannot be masculine, e.g. <i>die Au</i>	Köpcke (1982: 83)	---
VCCCC MONO→*F,*N	Monosyllabic nouns with the structure VCCCC cannot be feminine or neuter, e.g. <i>der Arzt</i>	Köpcke (1982: 83)	---
CVCCCC MONO→*F	Monosyllabic nouns with the structure CVCCCC cannot be feminine, e.g. <i>der Herbst</i>	Köpcke (1982: 83)	---
CCV MONO→*F	Monosyllabic nouns with the structure CCV cannot be feminine, e.g. <i>das Blei</i>	Köpcke (1982: 83)	---

CCVCCC MONO→*F,*N	Monosyllabic nouns with the structure CCVCCC cannot be feminine or neuter, e.g. <i>der Schrumpf</i>	Köpcke (1982: 83)	---
CCCV MONO→*M	Monosyllabic nouns with the structure CCCV cannot be masculine, e.g. <i>die Spreu</i>	Köpcke (1982: 83)	---
CCCVCCC MONO→*F,*N	Monosyllabic nouns with the structure CCCVCCC cannot be feminine or neuter, e.g. <i>der Strumpf</i>	Köpcke (1982: 83)	---
CCCVCC MONO→*F,*N	Monosyllables with the structure CCCVCC cannot be feminine or neuter, e.g. <i>der Strand</i>	Köpcke (1982: 83)	---

Appendix C: Sources for criterion 1

Constraint	Source
ACAD. DISCIPLINES→*M,*N	Wikipedia contributors. 29 April, 2016. Einzelwissenschaft. https://de.wikipedia.org/wiki/Einzelwissenschaft (10 July, 2016).
ALC. DRINKS→*F,*N	Wikipedia contributors. 15 September, 2015. Liste von Getränken. https://de.wikipedia.org/wiki/Liste_von_Getr%C3%A4nken (10 July, 2016).
ANG. SHAPES→*M,*F	Geometrische Formen und Muster. http://www.kinder-malvorlagen.com/zum-ausmalen/vorlagen-geometrische-formen.php (10 July, 2016). Rasch, Renate. 2016. Körper – Überblick. http://www.uni-landau.de/rasch/Grundlegende%20Geometrie/Vorlesungen/V7_Koerper_Ueberblick.pdf (10 July, 2016).
ANNOUNCE→*M,*F	Duden Synonymwörterbuch (2010): search terms <Ankündigung>, <Ansage>, <Urteil>, <Verkündigung>, <Vortrag>
BIRDS→*F,*N	Vögel. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 628-635.
BOATS→*F,*N	Wikipedia contributors. 4 February, 2016. Liste von Schiffstypen. https://de.wikipedia.org/wiki/Liste_von_Schiffstypen (10 July, 2016).
BODIES OF WATER→*N	Wikipedia contributors. 24 January, 2016. Gewässer. https://de.wikipedia.org/wiki/Kategorie:Liste_(Gew%C3%A4sser) (10 July, 2016).
CHEESES→*F,*N	Duden-Bibliothek Express (1993-2014): search term <Käse> Wikipedia contributors. 11 July, 2016. Liste von Käsesorten nach Herkunftsländern. https://de.wikipedia.org/wiki/Liste_von_K%C3%A4sesorten_nach_Herkunftslandern (18 July, 2016).
CHEMICAL COMP./SUBST.→*M,*F	Wikipedia contributors. 10 July, 2016. Grundchemikalie. https://de.wikipedia.org/wiki/Grundchemikalie (10 July, 2016). Wikipedia contributors. 3 July, 2016. Trivialname (Chemie). https://de.wikipedia.org/wiki/Trivialname_(Chemie) (10 July, 2016). Wikipedia contributors. 7 June, 2016. Chemikalienliste. https://de.wikipedia.org/wiki/Wikipedia:Redaktion_Chemie/Chemikalienliste (10 July, 2016).
CITIES/TOWNS→*M,*F	Herrmann, Franz. 2014. Die 308 Größten Städte der Welt. http://immobilien.trust-wi.de/2014/07/die-308-grossten-stadte-der-welt-2014-teil-1/ (10 July, 2016). Berger, Sabine. 2015. Die 10 schönsten Dörfer und Kleinstädte Europas. https://www.skyscanner.de/nachrichten/die-10-schoensten-doerfer-und-kleinstaedte-europas (10 July, 2016).
CONDUCT→*M,*N	Duden Synonymwörterbuch (2010): search terms <Moral>, <Ethik>
COUNTRIES→*M,*F	Alle Länder von A-Z. http://www.laenderdaten.de/laender.aspx (10 July, 2016).
DAYS→*F,*N	Feiertage 2016. http://www.schulferien.org/Feiertage/2016/feiertage_2016.html (10 July, 2016). Duden-Bibliothek Express (1993-2014): search term <Tag>
	Liste der Krankheiten von A-Z. http://www.welt.de/gesundheit/krankheiten-a-z/ (10 July, 2016).

DISEASES→*M,*N	Duden-Bibliothek Express (1993-2014): search term <Krankheit>
DOM. ANIMALS→*M,*F	Wikipedia contributors. 29 May, 2016. Liste domestizierter Tiere. https://de.wikipedia.org/wiki/Liste_domestizierter_Tiere (10 July, 2016).
DWELL→*M,*F	Duden Synonymwörterbuch (2010): search term <Heim>
ELAB. INSTR.→*M,*N	Wikipedia contributors. 11 July, 2016. Liste der Werkzeuge. https://de.wikipedia.org/wiki/Liste_der_Werkzeuge (13 July, 2016).
ELEMENTS→*M,*F	Royal Society of Chemistry. 2016. Periodic Table. www.rsc.org/periodic_table (10 July, 2016).
EXTROVERSION→*F,*N	Duden Synonymwörterbuch (2010): search terms <Hochmut>, <Mut>, <Mumm>, <Frechheit>, <Anmaßung>
F. HOLLOWES→*M,*F	Google Images: search terms <Dinge, die hohl sind>, <runde Dinge> (10 June, 2016).
FABRIC→*F,*N	Stoffkunde und Begriffserklärung. http://www.modeopfer110.de/mode-know-how/textillexikon-materialerklaerung.html (10 July, 2016).
FEAR→*M,*N	Duden Synonymwörterbuch (2010): search terms <Angst>, <Dringlichkeit>
FISH→*F,*N	Alle Fische von A-Z. http://www.lexikon-fische.de/ (10 July, 2016).
FLOWERS→*M,*N	Ziergarten. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 104-106.
FRUITS→*M,*N	Wikipedia contributors. 13 July, 2016. Liste der Obstarten. https://de.wikipedia.org/wiki/Liste_der_Obstarten (13 July, 2016).
GAMES→*M,*F	Wikipedia contributors. 22 June, 2016. Kartenspiel. https://de.wikipedia.org/wiki/Kartenspiel (10 July, 2016).
	Wikipedia contributors. 25 May, 2016. Brettspiel. https://de.wikipedia.org/wiki/Brettspiel (10 July, 2016).
	Alle Sportarten von A bis Z. https://www.decathlon.de/alle-sportarten-a-z.html (10 July, 2016).
GEN. PERSON→*F,*N; GEN. PERSON→*M,*F	Duden-Bibliothek Express (1993-2014): search terms <Person>, <jemand> and <Mensch>
GESTURES→*M,*N	Wikipedia contributors. 1 July, 2016. Liste von Gesten. https://de.wikipedia.org/wiki/Liste_von_Gesten (10 July, 2016).
GROUPS→*M,*N	Duden-Bibliothek Express (1993-2014): search terms <Gruppe>, <Gemeinde>, <Gemeinschaft>, <Gesellschaft>, <Mannschaft>
HUNTING→*M,*N	Duden-Bibliothek Express (1993-2014): search term <Jägersprache>
INSTITUTIONS→*M,*F	Wikipedia contributors. 3 July, 2016. Institution. https://de.wikipedia.org/wiki/Institution (10 July, 2016).
INTROVERSION→*M,*N	Duden Synonymwörterbuch (2010): search terms <Ängstlichkeit>, <Gehemmtheit>, <Sanftmut>, <Scheu>
IMPROPER SUPERORDINATES→*M,*F	---
INSECTS→*M,*N	Insekten. http://www.rohkostwiki.de/wiki/Insekten (10 July, 2016).
INTEGRATED PARTS→*F,*N	---
ISLANDS→*M,*F	Wikipedia contributors. 7 July, 2016. Liste der Inseln. https://de.wikipedia.org/wiki/Liste_der_Inseln_(nach_Staat) (10 July, 2016).
LANGUAGES→*M,*F	Top 100 languages by population. http://multilingualbooks.com/wp/soundandvision/2013/12/10/top-100-languages-population/ (10 July, 2016).
LEATHER TYPES→*M,*F	Lederarten. http://www.lederzentrum.de/wiki/index.php/Ledersorten (10 July, 2016).

LETTERS→*M,*F	Duden-Bibliothek Express (1993-2014): search term <Buchstabe>
LIVING/WORKING→*M,*F	Duden Synonymwörterbuch (2010): search term <Heim> Learn English free – job and work vocabulary. http://www.learnenglish.de/vocabulary/jobs.html (15 July, 2016).
MALE ANIMALS→*F,*N	Wikipedia contributors. 16 July, 2016. Liste der Bezeichnungen für Haus- und Wildtiere. https://de.wikipedia.org/wiki/Liste_der_Bezeichnung_f%C3%BCr_Haus-_und_Wildtiere (18 July, 2016).
MAMMALS→*F,*N	Säugetiere. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 640-644.
MEN→*F,*N	Duden-Bibliothek Express (1993-2014): search terms <Mann> and <Junge>
MESH→*M,*F	Duden Synonymwörterbuch (2010): search term <Netz>, <Masche>, <Geflecht>, <Gitter>
METALS→*M,*F	Royal Society of Chemistry. 2016. Periodic Table. < www.rsc.org/periodic_table > (10 July, 2016).
MON. UNITS→*F,*N	Currencies of the World. www.worldatlas.com/aatlas/infopage/currency (10 July, 2016).
MOUNTAINS→*F,*N	Wikipedia contributors. 30 June, 2016. Höchster Berg. https://de.wikipedia.org/wiki/H%C3%B6chster_Berg (10 July, 2016). Wikipedia contributors. 17 March, 2016. Liste der Berge oder Erhebungen in Europa. https://de.wikipedia.org/wiki/Liste_der_Berge_oder_Erhebungen_in_Europa (10 July, 2016). Wikipedia contributors. 25 June, 2016. Liste der höchsten Berge in Deutschland. https://de.wikipedia.org/wiki/Liste_der_h%C3%B6chsten_Berge_in_Deutschland (10 July, 2016).
MUS. COMP.→*M,*F	Types of Compositions. http://library.princeton.edu/departments/tsd/katmandu/music/mustyp.html (11 July, 2016).
MUS. INSTR.→*M,*F; MUS. INSTR.→*M,*N; MUS. INSTR.→*M	Musikinstrumente. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 560-568.
MUS. NOTES→*M,*F	Duden-Bibliothek Express (1993-2014): search term <Tonbezeichnung>
MUS. TERMS→*M,*F	Duden-Bibliothek Express (1993-2014): search term <Musik> Musiklexikon. http://www.inside-musik.de/html/musiklexikon.html (10 July, 2016).
MUSIC TYPES→*F,*N	Wikipedia contributors. 27 June, 2016. Stilrichtungen der Musik. https://de.wikipedia.org/wiki/Portal:Musik/Stilrichtungen_der_Musik (10 July, 2016).
NAME: MAN→*F,*N	Duden-Bibliothek Express (1993-2014): search term <Name>
NAME: WOMAN→*M,*N	Duden-Bibliothek Express (1993-2014): search term <Name>
NAUTICAL→*M,*N	Duden-Bibliothek Express (1993-2014): search term <Seemannssprache>
NO INFO.→*M,*F	Duden Synonymwörterbuch (2010): search term <Ding>
OPEN WATER→*M,*N	Wikipedia contributors. 24 January, 2016. Gewässer. https://de.wikipedia.org/wiki/Kategorie:Liste_(Gew%C3%A4sser) (10 July, 2016).
OPENINGS→*M,*F	---
ORGANS→*M,*F	Organe des Menschen. https://www.taschenhirn.de/mensch-und-natur/organe-des-menschen/ (10 July, 2016).

PCC GARMENT→*M,*F	Wikipedia contributors. 9 July, 2016. Liste der Kleidungsstücke. https://de.wikipedia.org/wiki/Liste_der_Kleidungsst%C3%BCcke (18 July, 2016).
PICTURES→*M,*F	Duden-Bibliothek Express (1993-2014): search term <Bild>
PLANTS→*F,*N	Zimmerpflanzen. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 104-107. Duden-Bibliothek Express (1993-2014): search term <Pflanze>
POWER/STRENGTH→*M,*N	Woxikon Synonymwörterbuch (2016): search term <Kraft>. http://synonyme.woxikon.de/synonyme/kraft.php (10 July, 2016).
PREP. DISHES→*M,*F	Duden-Bibliothek Express (1993-2014): search term <Gericht>, <Speise> Rezepte. http://www.essen-und-trinken.de/rezepte (18 July, 2016).
PRIMATES→*F,*N	Wikipedia contributors. 9 July, 2016. Primaten. https://de.wikipedia.org/wiki/Primaten (10 July, 2016).
REGIONS→*M,*F	Duden-Bibliothek Express (1993-2014): search terms <Landschaft>, <Region>, <Gebiet>, <Provinz>
RIVERS EUROPE→*M,*N	Wikipedia contributors. 4 January, 2016. Liste von Flüssen in Europa. https://de.wikipedia.org/wiki/Liste_von_Fl%C3%BCssen_in_Europa (10 July, 2016).
RIVERS NON-EUROPE→*F,*N	Wikipedia contributors. 6 July, 2016. Liste der längsten Flüsse der Erde. https://de.wikipedia.org/wiki/Liste_der_l%C3%A4ngsten_Fl%C3%BCsse_der_Erde (10 July, 2016).
RIVERS NON-GER.→*F,*N	Längste Flüsse der Welt. https://www.taschenhirn.de/geografie/laengste-fluesse-der-welt/ (10 July, 2016).
ROCKS/MINERALS →*F,*N	Duden-Bibliothek Express (1993-2014): search terms <Mineral> and <Gestein>
SCI. UNITS→*M,*F	International System of Units (SI). http://physics.nist.gov/cuu/Units/units.html (10 July, 2016).
SEASONS→*F,*N	Jahreszeiten – wann beginnen Frühling, Sommer, Herbst und Winter? http://www.wetter.de/cms/jahreszeiten-wann-beginnen-fruehling-sommer-herbst-und-winter-1832038.html (3 March, 2017).
SEMANTIC DIMINUTIVES→*M,*N	---
SEPARABLE PARTS→*M,*F	---
SHARP→*M,*N	Wikipedia contributors. 11 July, 2016. Liste der Werkzeuge. https://de.wikipedia.org/wiki/Liste_der_Werkzeuge (13 July, 2016).
SHEET/STRIP/SLAB→*M,*F	Google Images: search terms <Dinge, die flach sind>, <flache Objekte> (10 July, 2016).
SPEECH→*M,*N	Portz, Renate & Karvela, Ioanna. 2013. Studienführer zur Sprachpraxis. http://www.gs.uoa.gr/fileadmin/ga.uoa.gr/uploads/PDF/Odigos_PGA_2i_ekdosi.pdf (10 July, 2016).
STICK→*F,*N	Wikipedia contributors. 11 July, 2016. Liste der Werkzeuge. https://de.wikipedia.org/wiki/Liste_der_Werkzeuge (13 July, 2016).
SUP.→*M,*F	Internationale Waren- und Dienstleistungsklassifikation. Swiss Federal Institute of Intellectual Property. https://www.ige.ch/fileadmin/user_upload/Marken/d/nizza10_ob_d.pdf (10 July, 2016).
SUPERORDINATES OF INDIRECT REFERENCE→*M,*F	---
TEMPERATURES→*M,*N	Woxikon Synonymwörterbuch (2016): search terms <Hitze>, <Kälte>. http://synonyme.woxikon.de/synonyme/hitze.php ; http://synonyme.woxikon.de/synonyme/kaelte.php (10 July, 2016).

TIME: SHORT→*F,*N	Woxikon Synonymwörterbuch (2016): search term <kurze Zeit> http://synonyme.woxikon.de/synonyme/kurzezeit.php (11 July, 2016). Wahrig Synonymwörterbuch (Wissen.de, 2014-2016): search term <Augenblick> http://www.wissen.de/synonym/augenblick (11 July, 2016).
TIMES OF DAY→*F,*N	Swick, Edward 2016. How to say dates and times in German. http://www.dummies.com/how-to/content/how-to-say-dates-and-times-in-german.html (10 July, 2016).
TREES→*M,*N	Laubhölzer, Ziersträucher und Zierbäume. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 650-656.
UNITS OF TIME→*F,*N	Zeit. http://www.einheiten-umrechnen.org/zeit (10 July, 2016).
VEG.ROOT→*M,*N	Wikipedia contributors. 8 July, 2016. Wurzelgemüse. https://de.wikipedia.org/wiki/Wurzelgemüse (19 July, 2016).
WASTE→*F,*N	Wahrig Synonymwörterbuch (Wissen.de, 2014-2016): search term <Müll>. http://www.wissen.de/synonym/muell (11 July, 2016).
WOMEN→*M,*N	Duden-Bibliothek Express (1993-2014): search terms <Frau> and <Mädchen>
WORD CLASSES→*M,*F	Bopp, Stephan. 2000-2016. Einteilung der Wortklassen. http://www.canoo.net/services/OnlineGrammar/Wort/Ueberblick/Wortklasse.html (14 July, 2016).
YOUNG ANIMALS→*M,*F	Wikipedia contributors. 16 July, 2016. Liste der Bezeichnungen für Haus- und Wildtiere. https://de.wikipedia.org/wiki/Liste_der_Bezeichnung_f%C3%BCr_Haus-_und_Wildtiere (18 July, 2016).

Appendix D: Sources for criterion 4

Constraint	Examples of extra-linguistic evidence
ACAD. DISCIPLINES→*M,*N	Wikipedia contributors. 29 April, 2016. Einzelwissenschaft. https://de.wikipedia.org/wiki/Einzelwissenschaft (10 July, 2016).
ALC. DRINKS→*F,*N	Alkoholische Getränke. http://www.wissen.de/lexikon/alkoholische-getraenke (27 September, 2016).
ANG. SHAPES→*M,*F	Geometrische Formen und Muster. http://www.kinder-malvorlagen.com/zum-ausmalen/vorlagen-geometrische-formen.php (10 July, 2016).
BIRDS→*F,*N	Vögel. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 628-635.
BOATS→*F,*N	Wikipedia contributors. 4 February, 2016. Liste von Schiffstypen. https://de.wikipedia.org/wiki/Liste_von_Schiffstypen (10 July, 2016).
BODIES OF WATER→*N	Wikipedia contributors. 24 January, 2016. Gewässer. https://de.wikipedia.org/wiki/Kategorie:Liste_(Gew%C3%A4sser) (10 July, 2016).
CHEESES→*F,*N	Wikipedia contributors. 11 July, 2016. Liste von Käsesorten nach Herkunftsländern. https://de.wikipedia.org/wiki/Liste_von_K%C3%A4sesorten_nach_Herkunftsl%C3%A4ndern (18 July, 2016).
CHEMICAL COMP./SUBST.→*M,*F	Wikipedia contributors. 7 July, 2016. Chemikalienliste. https://de.wikipedia.org/wiki/Wikipedia:Redaktion_Chemie/Chemikalienliste (10 July, 2016).
CITIES/TOWNS→*M,*F	Wikipedia contributors. 19 September, 2016. Liste der Städte in Deutschland. https://de.wikipedia.org/wiki/Liste_der_St%C3%A4dte_in_Deutschland (26 September, 2016).
COUNTRIES→*M,*F	Alle Länder von A-Z. http://www.laenderdaten.de/laender.aspx (10 July, 2016).
DAYS→*F,*N	Tage der Woche. http://german.about.com/library/blatein02.htm (26 September, 2016). Feiertage 2016. http://www.schulferien.org/Feiertage/2016/feiertage_2016.html (10 July, 2016).
DISEASES→*M,*N	Liste der Krankheiten von A-Z. http://www.welt.de/gesundheit/krankheiten-a-z/ (27 September, 2016).
DOM. ANIMALS→*M,*F	Haustiere. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 141.
DWELL→*M,*F	Wohnungsarten. https://welcome.dresden.de/de/wohnen/arten.php (29 September, 2016).
ELEMENTS→*M,*F	Chemische Elemente. http://www.wissen.de/lexikon/chemische-elemente (27 September, 2016).
FABRIC→*F,*N	Stoffkunde und Begrifferklärung. http://www.modeopfer110.de/mode-know-how/textillexikon-materialerklaerung.html (10 July, 2016).
FISH→*F,*N	Alle Fische von A-Z. http://www.lexikon-fische.de/ (26 September, 2016).
FLOWERS→*M,*N	Gartenblumen. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 118-119.
FRUITS→*M,*N	Wikipedia contributors. 13 July, 2016. Liste der Obstarten. https://de.wikipedia.org/wiki/Liste_der_Obstarten (13 July, 2016).
GAMES→*M,*F	Freizeit, Spiel, Sport. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 474-531.
GEN. PERSON→*F,*N;	Berufsbezeichnungen. http://www.bewerben.com/content/berufsbezeichnungen (29 September, 2016).

GEN. PERSON→*M,*F	
GESTURES→*M,*N	Wikipedia contributors. 1 July, 2016. Liste von Gesten. https://de.wikipedia.org/wiki/Liste_von_Gesten (10 July, 2016).
GROUPS→*M,*N	Substantive und Artikel: Gruppenbezeichnungen. Duden – <i>Wissen – Üben – Testen: Englisch 7. Klasse</i> . 4 th ed. 2016. Berlin: Bibliographisches Institut. 6.
HUNTING→*M,*N	Jagd. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 164-165.
INSECTS→*M,*N	Insekten. http://www.rohkostwiki.de/wiki/Insekten (10 July, 2016).
INSTITUTIONS→*M,*F	Wikipedia contributors. 3 July, 2016. Institution, https://de.wikipedia.org/wiki/Institution (10 July, 2016).
ISLANDS→*M,*F	Wikipedia contributors. 7 July, 2016. Liste der Insellisten. https://de.wikipedia.org/wiki/Liste_der_Insellisten_(nach_Staat) (10 July, 2016).
LANGUAGES→*M,*F	Sprachen der Welt: Alle Sprachen von A-Z. http://www.spracheninfos.de/a-z.aspx (30 September, 2016).
LEATHER TYPES→*M,*F	Lederarten. http://www.lederzentrum.de/wiki/index.php/Ledersorten (10 July, 2016).
LETTERS→*M,*F	Alphabet. http://www.wissen.de/lexikon/alphabet (27 September, 2016).
LIVING/WORKING→*M,*F	Wohnungsarten. https://welcome.dresden.de/de/wohnen/arten.php (29 September, 2016).
MALE ANIMALS→*F,*N	Wikipedia contributors. 16 July, 2016. Liste der Bezeichnungen für Haus- und Wildtiere: Männchen. https://de.wikipedia.org/wiki/Liste_der_Bezeichnung_f%C3%BCr_Haus-_und_Wildtiere (18 July, 2016).
MAMMALS→*F,*N	Mammals/Säugetiere. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 640-645.
MEN→*F,*N	Wikipedia contributors. 7 September, 2014. Mann: Bezeichnung. https://de.wikipedia.org/wiki/Kategorie:Mann_(Bezeichnung) (29 September, 2016).
METALS→*M,*F	Metalle. http://www.chemie.de/lexikon/Metalle.html (27 September, 2016).
MON. UNITS→*F,*N	Liste der Währungen aller Länder. http://www.laenderdaten.de/wirtschaft/waehrungen.aspx (30 September, 2016).
MOUNTAINS→*F,*N	Wikipedia contributors. 17 March, 2016. Liste der Berge oder Erhebungen in Europa. https://de.wikipedia.org/wiki/Liste_der_Berge_oder_Erhebungen_in_Europa (10 July, 2016).
MUS. INSTR.→*M,*F; MUS. INSTR.→*M,*N; MUS. INSTR.→*M	Musikinstrumente. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 560-579.
MUS. COMP.→*M,*F	Types of compositions. http://library.princeton.edu/departments/tsd/katmandu/music/mustyp.html (11 July, 2016).
MUS. NOTES→*M,*F	Musiknotation. The Oxford-Duden Pictorial German Dictionary, 2nd edn. 1994. Oxford: Oxford University Press. 556-559.
MUS. TERMS→*M,*F	Musiklexikon. http://www.inside-musik.de/html/musiklexikon.html (10 July, 2016).
MUSIC TYPES→*F,*N	Wikipedia contributors. 27 June, 2016. Stilrichtungen der Musik. https://de.wikipedia.org/wiki/Portal:Musik/Stilrichtungen_der_Musik (10 July, 2016).
NAME: MAN→*F,*N	Wikipedia contributors. 4 August, 2016. Männlicher Vorname. https://de.wikipedia.org/wiki/Kategorie:Männlicher_Vorname (29 September, 2016).

NAME: WOMAN→*M,*N	Wikipedia contributors. 17 September, 2016. Weiblicher Vorname. https://de.wikipedia.org/wiki/Kategorie:Weiblicher_Vorname (29 September, 2016).
NAUTICAL→*M,*N	Die Seemannssprache verstehen lernen. http://seemannssprache.com/ (29 September, 2016).
ORGANS→*M,*F	Organe des Menschen. https://www.taschenhirn.de/mensch-und-natur/organe-des-menschen/ (10 July, 2016).
PICTURES→*M,*F	Visuelle Medien. http://www.didactics.eu/index.php?id=804 (29 September, 2016).
PLANTS→*F,*N	Launer, Annette. 2016. Pflanzen-Liste nach deutschen Namen. http://www.pflanzenliebe.de/innen/innen_liste_deutsch.html (4 October, 2016).
PREP. DISHES→*M,*F	Rezepte. http://www.essen-und-trinken.de/rezepte (18 July, 2016).
PRIMATES→*F,*N	Wikipedia contributors. 9 July, 2016. Primaten. https://de.wikipedia.org/wiki/Primaten (10 July, 2016).
REGIONS→*M,*F	Wikipedia contributors. 2 April, 2013. Deutschland nach Region. https://de.wikipedia.org/wiki/Kategorie:Deutschland_nach_Region (29 September, 2016).
RIVERS EUROPE→*M,*N	Wikipedia contributors. 4 January, 2016. Liste von Flüssen in Europa. https://de.wikipedia.org/wiki/Liste_von_Fl%C3%BCssen_in_Europa (10 July, 2016).
RIVERS NON-EUROPE→*F,*N	Flüsse außerhalb Europas. http://www.go7seas-kreuzfahrten.de/ausserhalb-europas (29 September, 2016).
RIVERS NON-GER.→*F,*N	Längste Flüsse der Welt. https://www.taschenhirn.de/geografie/laengste-fluesse-der-welt/ (10 July, 2016).
ROCKS/MINERALS →*F,*N	Wikipedia contributors. 2 May, 2014. Minerale und Gesteine. https://de.wiktionary.org/wiki/Verzeichnis:Deutsch/Minerale_und_Gesteine (29 September, 2016).
SCI. UNITS→*M,*F	SI Einheiten. http://www.wissen.de/lexikon/si-einheiten (27 September, 2016).
SEASONS→*F,*N	Jahreszeiten – wann beginnen Frühling, Sommer, Herbst und Winter? http://www.wetter.de/cms/jahreszeiten-wann-beginnen-fruehling-sommer-herbst-und-winter-1832038.html (3 March, 2017).
SHARP→*M,*N	Verbotene Gegenstände: Spitze/scharfe Waffen und scharfe Objekte. Bundesministerium für Verkehr, Innovation und Technologie, November 2006. https://www.tui.at/fileadmin/redaktion/redaktion/PDFs/Diverses/verbotene_gegenstaende.pdf (27 September, 2016).
STICK→*F,*N	Malik, Miriam. 2015. Selbstverteidigung mit Alltagsgegenständen: Stockähnliche Gegenstände. http://miriam-malik.de/selbstverteidigung-mit-alltagsgegenstaenden/ (30 September, 2016).
SUP.→*M,*F	Oberbegriffe. https://www.grundschulmaterial.de/medien/Deutsch/Klasse%203/Grammatik/Nomen/Oberbegriffe%20suchen/id/100234/ (27 September, 2016).
TEMPERATURES→*M,*N	Wikipedia contributors. 8 February, 2016. Temperatur. https://de.wikipedia.org/wiki/Temperatur (10 July, 2016).
TIMES OF DAY→*F,*N	Swick, Edward. 2016. How to say dates and times in German. http://www.dummies.com/how-to/content/how-to-say-dates-and-times-in-german.html (10 July, 2016).
TREES→*M,*N	Baumarten von A-Z nach deutschen Namen. http://www.baumkunde.de/baumlisten/baumliste_az.php (4 October, 2016).
UNITS OF TIME→*F,*N	Zeit. http://www.einheiten-umrechnen.org/zeit (10 July, 2016).

VEG.ROOT→*M,*N	Wikipedia contributors. 8 July, 2016. Wurzelgemüse. https://de.wikipedia.org/wiki/Wurzelgemüse (19 July, 2016).
WASTE→*F,*N	Wikipedia contributors. 6 June, 2016. Abfall. https://de.wikipedia.org/wiki/Abfall (10 July, 2016).
WOMEN→*M,*N	Wikipedia contributors. 7 September, 2014. Frau: Bezeichnung. https://de.wikipedia.org/wiki/Kategorie:Frau_(Bezeichnung) (29 September, 2016).
WORD CLASSES→*M,*F	Bopp, Stephan. 2000-2016. Einteilung der Wortklassen. http://www.canoo.net/services/OnlineGrammar/Wort/Ueberblick/Wortklasse.html (14 July, 2016).
YOUNG ANIMALS→*M,*F	Wikipedia contributors. 16 July, 2016. Liste der Bezeichnungen für Haus- und Wildtiere: Junges. https://de.wikipedia.org/wiki/Liste_der_Bezeichnung_f%C3%BCr_Haus-_und_Wildtiere (18 July, 2016).

Appendix E: Summary of results from Chapter 2

Constraint	1. Share of candidate nouns	2. Parallels in other languages	3. Parallels within German language	4. Cultural relevance and clarity (semantic)	5. Ability to account for exceptions	6. Diachronic productivity	TOTAL
ELAB. INSTR. → *M, *N	X	X	X	X	X	-	0
EXTROVERSION → *F, *N	X	X	X	X	X	-	0
PCC GARMENT → *M, *F	X	X	X	X	X	X	0
ANNOUNCE → *M, *F	X	X	X	X	X	-	0
-IV UNSTRESSED → *F, *N	X	X	X	n/a	X	-	0
-EKT → *M, *F	X	X	X	n/a	X	-	0
TIME: SHORT → *F, *N	X	X	X	X	X	-	0
OPEN WATER → *M, *N	X	X	X	X	X	-	0
SEMANTIC DIMINUTIVES → *M, *N	X	X	X	X	X	-	0
SEPARABLE PARTS → *M, *F	X	X	X	X	X	-	0
SUPERORDINATES OF INDIRECT REFERENCE → *M, *F	X	X	X	X	X	-	0
INTEGRATED PARTS → *F, *N	X	X	X	X	X	-	0
IMPROPER SUPERORDINATES → *M, *F	X	X	X	X	X	-	0
PREP. DISHES → *M, *F	X	X	X	✓	X	X	1
SHEET/STRIP/SLAB → *M, *F	X	✓	X	X	-	X	1
INTROVERSION → *M, *N	✓	X	X	X	X	-	1
BOATS → *F, *N	X	X	X	✓	X	-	1
HUNTING → *M, *N	X	X	X	✓	X	-	1
NAUTICAL → *M, *N	X	X	X	✓	X	X	1
OPENINGS → *M, *F	X	✓	X	X	X	X	1
GESTURES → *M, *N	X	X	X	✓	X	-	1
POWER/STRENGTH → *M, *N	✓	X	X	X	X	-	1
SHARP → *M, *N	X	X	X	✓	X	-	1
SPEECH → *M, *N	✓	X	X	X	X	-	1
-ER → *M, *F	X	X	✓	n/a	X	X	1
-ER → *M, *N	X	X	✓	n/a	X	X	1
-LT MONO → *M, *F	X	X	X	n/a	X	✓	1
-[I]C MONO → *F, *N	X	X	X	n/a	X	✓	1
VER- → *F	X	✓	X	n/a	X	-	1
-IER /i:r/ POLY. → *M, *F	X	✓	X	n/a	X	-	1
-AR SUFF. → *F, *N	X	X	✓	n/a	X	-	1
-AR SUFF. → *M, *F	X	X	✓	n/a	X	-	1
-EE/-ÉE/-É → *M, *F	X	X	✓	n/a	X	-	1
-EE/-ÉE/-É → *F, *N	X	X	✓	n/a	X	-	1
-EE/-ÉE/-É → *M, *N	X	X	✓	n/a	X	-	1
-EHR/-ÄHR → *M, *N	X	X	X	n/a	✓	-	1
-SAL → *M, *N	X	X	✓	n/a	X	-	1
DISEASES → *M, *N	X	X	X	✓	✓	X	2
MUS. TERMS → *M, *F	X	X	X	✓	✓	-	2
MAMMALS → *F, *N	X	X	X	✓	✓	-	2
MUS. INSTR. → *M, *F	X	✓	X	✓	X	X	2
MUS. INSTR. → *M, *N	X	✓	X	✓	X	-	2
GROUPS → *M, *F	X	X	✓	✓	X	X	2
ORGANS → *M, *F	X	✓	X	✓	X	-	2
F. HOLLOW → *M, *F	X	✓	X	X	✓	X	2
CONDUCT → *M, *N	✓	X	X	X	✓	-	2
MESH → *M, *F	✓	X	X	X	✓	-	2
-AT → *F, *N	X	✓	✓	n/a	X	X	2
-IV → *F, *N	X	✓	✓	n/a	X	-	2
-NIS → *M, *N	X	✓	✓	n/a	X	X	2
-O → *M, *F	X	✓	✓	n/a	X	X	2

-TUM→*F,*N	X	✓	✓	n/a	X	X	2
-IT→*M,*F	X	X	✓	n/a	✓	-	2
-ON→*M,*F	X	X	✓	n/a	X	✓	2
-ST/-SCHT/-ZT→*F,*N	X	X	✓	n/a	X	✓	2
-ENT→*M,*F	X	✓	✓	n/a	X	-	2
-AL→*M,*F	X	✓	X	n/a	X	✓	2
-EL→*M,*F	X	X	✓	n/a	✓	X	2
-EL→*F,*N	X	X	✓	n/a	✓	-	2
-EL→*M,*N	X	X	✓	n/a	✓	-	2
LIVING/WORKING→*M,*F	X	X	X	X	✓	✓	2
RIVERS EUROPE→*M,*N	X	✓	X	✓	X	-	2
MUS. COMP.→*M,*F	X	X	X	✓	✓	-	2
PICTURES→*M,*F	X	X	X	✓	✓	✓	3
DOM. ANIMALS→*M,*F	X	X	X	✓	✓	✓	3
PLANTS→*F,*N	X	✓	X	✓	✓	X	3
BIRDS→*F,*N	X	✓	X	✓	✓	-	3
STICK→*F,*N	X	✓	X	✓	X	✓	3
UNITS OF TIME→*F,*N	X	✓	✓	✓	X	-	3
GEN. PERSON→*M,*F	X	✓	✓	✓	X	X	3
FEAR→*M,*N	✓	X	X	X	✓	✓	3
NO INFO.→*M,*F	✓	X	X	X	✓	✓	3
-EN→*F,*N	X	X	✓	n/a	✓	✓	3
-EN→*M,*F	X	X	✓	n/a	✓	✓	3
-FT/-CHT→*M,*N	X	✓	X	n/a	✓	✓	3
MONOSYLL.→*F,*N	X	X	✓	n/a	✓	✓	3
-NIS→*M,*F	X	✓	✓	n/a	X	✓	3
DWELL→*M,*F	X	✓	X	✓	✓	✓	4
INSTITUTIONS→*M,*F	X	✓	✓	✓	✓	✓	5
-RIS→*M,*N	✓	X	X	n/a	X	-	1
-EKT→*F,*N	✓	X	X	n/a	X	X	1
-/tj/;/Ntj/;/Nj// MONO→*F,*N	✓	X	X	n/a	X	-	1
KR- MONO→*F,*N	✓	X	X	n/a	X	-	1
TR→*F,*N	✓	X	X	n/a	-	-	1
VCC MONO→*F,*N	✓	X	X	n/a	X	-	1
-IG→*F,*N	✓	X	X	n/a	X	-	1
ANG. SHAPES→*M,*F	✓	X	X	✓	X	-	2
TEMPERATURES→*M,*N	✓	X	X	✓	X	-	2
BODIES OF WATER→*N	✓	X	X	✓	X	-	2
D- MONO→*F	✓	X	X	n/a	X	✓	2
-AR→*M,*F	✓	X	X	n/a	X	✓	2
CCCVC MONO→*F,*N	✓	X	X	n/a	X	✓	2
-IG/-ICH→*F,*N	✓	X	X	n/a	X	✓	2
-EKT→*F	✓	X	X	n/a	X	✓	2
-AN→*M,*F	✓	✓	X	n/a	X	-	2
-EDER→*M,*F	✓	✓	X	n/a	X	-	2
-MA→*M,*F	✓	✓	X	n/a	X	-	2
-IVE→*M,*N	✓	✓	X	n/a	X	-	2
-INER→*F,*N	✓	✓	X	n/a	X	-	2
V _{STRESSED} [l]→*M,*F	✓	✓	X	n/a	X	-	2
V _{STRESSED} [m]→*M,*F	✓	✓	X	n/a	X	-	2
-ETTE→*M,*N	✓	X	✓	n/a	X	-	2
-IE [iə]→*M,*N	✓	X	✓	n/a	X	-	2
-IN [i:n]→*M,*F	✓	X	✓	n/a	X	-	2
-IT→*F,*N	✓	X	✓	n/a	X	-	2
-ON UNSTRESS.→*M,*F	✓	X	✓	n/a	X	-	2
-EUR/ÖR→*F,*N	✓	X	✓	n/a	X	-	2
-EUR/ÖR→*N	✓	X	✓	n/a	X	-	2
VCCC MONO→*M	✓	X	X	n/a	✓	-	2
DR-→*F,*N	✓	X	X	n/a	✓	-	2
-ANG→*F,*N	✓	X	X	n/a	✓	-	2

∅V _{LONG} MONO→*F	✓	X	X	n/a	✓	-	2
-ILLE→*M,*N	✓	X	X	n/a	✓	-	2
-ITIS→*M,*N	✓	✓	X	n/a	X	-	2
-UR/-UHR /u:r/→*M,*N	✓	X	X	n/a	✓	-	2
VC MONO→*M,*F	✓	X	X	n/a	✓	-	2
CV MONO→*N	✓	X	X	n/a	X	✓	2
CVCCC MONO→*F,*N	✓	X	X	n/a	X	✓	2
CCVCC→*F,*N	✓	X	X	n/a	X	✓	2
-B→*F	✓	X	X	n/a	X	✓	2
R- MONO→*F	✓	X	X	n/a	✓	✓	3
-IV STRESSED→*M,*F	✓	X	✓	n/a	✓	-	3
-ADE→*M,*N	✓	✓	✓	n/a	X	-	3
-ANER→*F,*N	✓	✓	✓	n/a	X	-	3
-ANT→*F,*N	✓	✓	✓	n/a	X	-	3
-ANZ→*M,*N	✓	✓	✓	n/a	X	-	3
-INE→*M,*N	✓	✓	✓	n/a	X	-	3
-IST→*F,*N	✓	✓	✓	n/a	X	-	3
-IUM→*M,*F	✓	✓	✓	n/a	X	-	3
-SKOP→*M,*F	✓	✓	✓	n/a	X	-	3
-(C)(C)// MONO →*F,*N	✓	X	✓	n/a	X	✓	3
-ST/-SCHT/-ZT→*N	✓	X	✓	n/a	X	✓	3
DEVERBAL -T→*M,*N	✓	X	✓	n/a	✓	-	3
-IS→*M,*N	✓	X	✓	n/a	✓	-	3
-IT→*F	✓	X	✓	n/a	✓	-	3
V _{LONG} C MONO→*F	✓	✓	X	n/a	✓	X	3
VER- NO SUFF. → *F	✓	✓	X	n/a	X	✓	3
-L MONO→*F	✓	X	X	n/a	✓	✓	3
-N(C)(C) MONO→*F,*N	✓	X	X	n/a	✓	✓	3
[[C]- MONO→*F,*N	✓	X	X	n/a	✓	✓	3
V _{LONG} MONO→*F	✓	X	X	n/a	✓	✓	3
-ETT→*M,*F	✓	X	X	n/a	✓	✓	3
T- MONO→*F	✓	X	X	n/a	✓	✓	3
V _{SHORT} STOP C→*F,*N	✓	X	X	n/a	✓	✓	3
VV _{DIPHTH} C MONO→*F	✓	X	X	n/a	✓	✓	3
CCVV _{DIPHTHONG} - MONO→*F	✓	X	X	n/a	✓	✓	3
CVC MONO→*F	✓	X	X	n/a	✓	✓	3
CVCC MONO→*F	✓	X	X	n/a	✓	✓	3
CVCC MONO→*N	✓	X	X	n/a	✓	✓	3
-ÜR/-ÜHR /y:r/→*M,*N	✓	X	X	n/a	✓	✓	3
-OR→*F,*N	✓	X	✓	n/a	X	✓	3
-SAL→*M	✓	X	✓	n/a	X	✓	3
-THEK→*M,*N	✓	✓	✓	n/a	X	-	3
-UR→*M,*N	✓	✓	✓	n/a	X	-	3
-ENT→*F,*N	✓	✓	✓	n/a	X	-	3
-ISMUS→*F,*N	✓	✓	✓	n/a	X	-	3
-US→*F,*N	✓	X	✓	n/a	✓	-	3
-/r/ STOP (C) MONO→*F,*N	✓	X	X	n/a	✓	✓	3
CCVC MONO→*F,*N	✓	X	X	n/a	✓	✓	3
-SAL→*M,*F	✓	X	✓	n/a	X	✓	3
ACAD. DISCIPLINES→*M,*N	✓	X	X	✓	✓	-	3
PRIMATES→*F,*N	✓	X	X	✓	✓	-	3
CHEESES→*F,*N	✓	X	X	✓	✓	-	3
INSECTS→*M,*N	✓	X	X	✓	✓	-	3
ISLANDS→*M,*F	✓	✓	X	✓	X	-	3
VEG.ROOT→*M,*N	✓	✓	X	✓	X	-	3
MUSIC TYPES→*F,*N	✓	X	X	✓	X	✓	3
RIVERS NON-EUROPE→*F,*N	✓	✓	X	✓	X	-	3
WASTE→*F,*N	✓	X	X	✓	X	✓	3
WORD CLASSES→*M,*F	✓	✓	X	✓	X	-	3
RIVERS NON-GER.→*F,*N	✓	✓	X	✓	X	-	3

MUS. NOTES→*M,*F	✓	✓	x	✓	x	-	3
TIMES OF DAY→*F,*N	✓	x	✓	✓	x	-	3
MOUNTAINS→*F,*N	✓	✓	x	✓	x	-	3
SEASONS→*F,*N	✓	✓	x	✓	x	-	3
-ING→*M,*F	✓	✓	✓	n/a	x	✓	4
-ION→*M,*N	✓	✓	✓	n/a	x	✓	4
-ENZ→*M,*N	✓	✓	✓	n/a	x	✓	4
-AGE→*M,*N	✓	✓	✓	n/a	x	✓	4
-IK→*M,*N	✓	✓	✓	n/a	x	✓	4
-TÄT→*M,*N	✓	✓	✓	n/a	x	✓	4
-TUM→*M,*F	✓	✓	✓	n/a	x	✓	4
-TUM→*F	✓	✓	✓	n/a	x	✓	4
-NIS→*M	✓	✓	✓	n/a	x	✓	4
-MENT→*M,*F	✓	✓	✓	n/a	x	✓	4
-EN→*F	✓	x	✓	n/a	✓	✓	4
-ER→*F,*N	✓	x	✓	n/a	✓	✓	4
-C S MONO →*F,*N	✓	x	✓	n/a	✓	✓	4
-A→*M,*N	✓	✓	✓	n/a	✓	x	4
-IV→*M,*F	✓	✓	✓	n/a	✓	-	4
-IV→*F	✓	✓	✓	n/a	✓	-	4
-SCHAFT→*M,*N	✓	✓	✓	n/a	✓	-	4
-(O)LOGE→*F,*N	✓	✓	✓	n/a	✓	-	4
LETTERS→*M,*F	✓	✓	✓	✓	x	-	4
MALE ANIMALS→*F,*N	✓	✓	✓	✓	x	-	4
REGIONS→*M,*F	✓	✓	✓	✓	x	-	4
ELEMENTS→*M,*F	✓	✓	x	✓	✓	-	4
YOUNG ANIMALS→*M,*F	✓	✓	x	✓	✓	-	4
FISH→*F,*N	✓	x	x	✓	✓	✓	4
LEATHER TYPES→*M,*F	✓	✓	x	✓	✓	-	4
TREES→*M,*N	✓	✓	x	✓	x	✓	4
MUS. INSTR. →*M	✓	✓	x	✓	x	✓	4
FRUITS→*M,*N	✓	✓	x	✓	x	✓	4
SCI. UNITS→*M,*F	✓	x	✓	✓	✓	-	4
MON. UNITS→*F,*N	✓	x	✓	✓	x	✓	4
ALC. DRINKS→*F,*N	✓	✓	x	✓	✓	✓	5
CITIES/TOWNS→*M,*F	✓	✓	✓	✓	✓	-	5
COUNTRIES→*M,*F	✓	✓	✓	✓	✓	-	5
LANGUAGES→*M,*F	✓	✓	✓	✓	✓	-	5
ROCKS/MINERALS →*F,*N	✓	✓	✓	✓	✓	-	5
FABRIC→*F,*N	✓	✓	✓	✓	✓	-	5
NAME: MAN→*F,*N	✓	✓	✓	✓	✓	-	5
NAME: WOMAN→*M,*N	✓	✓	✓	✓	✓	-	5
CHEMICAL COMP./SUBST. →*M,*F	✓	✓	x	✓	✓	✓	5
DAYS→*F,*N	✓	✓	✓	✓	x	✓	5
GAMES→*M,*F	✓	✓	x	✓	✓	✓	5
SUP. →*M,*F	✓	✓	x	✓	✓	✓	5
FLOWERS→*M,*N	✓	✓	x	✓	✓	✓	5
METALS→*M,*F	✓	✓	x	✓	✓	✓	5
GROUPS→*M,*N	✓	x	✓	✓	✓	✓	5
-CHEN→*M,*F	✓	✓	✓	n/a	✓	✓	5
CONVERSION→*M,*F	✓	✓	✓	n/a	✓	✓	5
-EI→*M,*N	✓	✓	✓	n/a	✓	✓	5
-ER SUFF. →*F,*N	✓	✓	✓	n/a	✓	✓	5
-HEIT/-(IG)KEIT→*M,*N	✓	✓	✓	n/a	✓	✓	5
-IE ['i:]→*M,*N	✓	✓	✓	n/a	✓	✓	5
-IN SUFF. →*M,*N	✓	✓	✓	n/a	✓	✓	5
-UNG→*M,*N	✓	✓	✓	n/a	✓	✓	5
-LING→*F,*N	✓	✓	✓	n/a	✓	✓	5
-UM→*M,*F	✓	✓	✓	n/a	✓	✓	5
DEVERBAL STEM→*F,*N	✓	✓	✓	n/a	✓	✓	5

-AT→*M,*F	✓	✓	✓	n/a	✓	✓	5
-E.SUFFIX→*M,*N	✓	✓	✓	n/a	✓	✓	5
-E→*M,*N	✓	✓	✓	n/a	✓	✓	5
GE-→*M,*F	✓	✓	✓	n/a	✓	✓	5
GEN. PERSON→*F,*N	✓	✓	✓	✓	✓	✓	6
WOMEN→*M,*N	✓	✓	✓	✓	✓	✓	6
MEN→*F,*N	✓	✓	✓	✓	✓	✓	6

Appendix F: Summary of results from Chapter 3

Noun	Constraints	Head noun	Constraints	OGAT	OGAT II
<i>das Abgangszeugnis</i>		<i>Zeugnis</i>	-NIS→*M	X	X
<i>das Allegro</i>	-O→*F MUS. TERM/COMP→*M			✓	✓
<i>das Anachoretentum</i>	-TUM→*M,*F			✓	✓
<i>das Anderkonto</i>		<i>Konto</i>	-O→*F SEM. EQUIV.→*M,*N (DIE BANKVERBINDUNG)	X	X
<i>das Antependium</i>	-IUM→*M,*F			✓	✓
<i>das Arrival</i>	-AL ENG.→*M,*F SEM. EQUIV.→*M,*N (DIE ANKUNFT)			X	✓
<i>das Bakschisch</i>	SEM. EQUIV.→*M,*F (DAS TRINGELD)			✓	✓
<i>das Benediktenkraut</i>	PLANTS→*N FLOWERS→*M,*N	<i>Kraut</i>	CCVC MONO→*F,*N CCVVDIPHTE MONO→*F VV _{DIPHTE} MONO→*F SUP.→*M,*F ¹	X	✓
<i>das Betäubungsmittel</i>	CHEM. COMP./SUBST.→*M,*F	<i>Mittel</i>	SUP.→*M,*F	✓	✓
<i>das Blackjack</i>	GAMES→*M,*F SEM. EQUIV.→*M,*F (DAS SIEBZEHNUNDEVIER)			✓	✓
<i>das Bootcamp</i>	SEM. EQUIV.→*M,*F (DAS TRAININGSLAGER)	<i>Camp</i>	DWELL→*M -N(C)(C)→*F,*N CVCC MONO→*F SEM. EQUIV.→*M,*F (DAS LAGER)	✓	✓
<i>das Coaching</i>	-ING→*M,*F SEM. EQUIV.→*M,*N (DIE BETREUUNG)			X	✓
<i>das Dazutun</i>	CONVERSION→*M,*F			✓	✓
<i>das Deleatur</i>	SYMBOLS→*M,*F			✓	✓
<i>das Derivat</i>	-AT→*M,*F			✓	✓
<i>das Deutschsprechen</i>	CONVERSION→*M,*F -EN→*F			✓	✓
<i>das Diaphanbild</i>	PICTURES→*M	<i>Bild</i>	CVCC MONO→*F -[I]C MONO→*F PICTURES→*M	✓	✓
<i>das Drogengeschäft</i>	GE→*M,*F	<i>Geschäft</i>	GE→*M,*F	✓	✓
<i>das Eclair</i>	SEM. EQUIV.→*F,*N (DER LIEBESKNOCHEN) -AIR [ɛ:p]→*M,*F			X	✓
<i>das Eigentumsrecht</i>		<i>Recht</i>	R- MONO→*F -FT/-CHT→*M CVCC MONO→*F	✓	✓
<i>das Epistyl</i>	V _{STRESSED} [I]→*M,*F			✓	✓
<i>das Erschöpfungssyndrom</i>	V _{STRESSED} [m]→*M,*F (DISEASES→*M/DISEASES →*N)	<i>Syndrom</i>	V _{STRESSED} [m]→*M,*F	✓	✓
<i>das Fahrenheit</i>	SCI. UNITS→*M,*F			✓	✓
<i>das Farsi</i>	LANGUAGES→*M,*F			✓	✓

¹ Justified in Zubin & Köpcke (1986: 157f.)

<i>das Ferrum</i>	METALS→*M,*F -UM→*M,*F ELEMENTS→*M,*F			✓	✓
<i>das Filet</i>	-ET ['e:]→*M,*F SEM. EQUIV.→*M,*F (DAS STEAK)			✓	✓
<i>das Flettnerruder</i>	-ER→*F,*N	<i>Ruder</i>	-ER→*F,*N	X	X
<i>das Freixemplar</i>	-AR→*M,*F	<i>Exemplar</i>	-AR→*M,*F	✓	✓
<i>das Fungizid</i>	CHEM. COMP./SUBST.→*M,*F			✓	✓
<i>das Gebärdenspiel</i>		<i>Spiel</i>	-L MONO→*F [[]C- MONO→*F,*N V _{LONG} (C) MONO→*F CCVC MONO→*F,*N SUP.→*M,*F ² DEVERBAL STEM→*F,*N	X	X
<i>das Gegenüber</i>	CONVERSION→*M,*F -ER→*F,*N GEN. PERSON→*F,*N			X	✓
<i>das Geleucht</i>	GE-→*M,*F			✓	✓
<i>das Getändel</i>	GE-→*M,*F			✓	✓
<i>das Goldhähnchen</i>	-CHEN→*M,*F BIRDS→*F	<i>Hähnchen</i>	-CHEN→*M,*F BIRDS→*F	✓	✓
<i>das Haben</i>	CONVERSION→*M,*F -EN→*F			✓	✓
<i>das Hallenbad</i>		<i>Bad</i>	V _{LONG} (C) MONO→*F CVC MONO→*F	X	X
<i>das Hexaeder</i>	-EDER→*M,*F -ER→*F,*N			X	✓
<i>das His</i>	MUS. NOTES→*M,*F MUS. TERM/COMP→*M CVC MONO→*F			✓	✓
<i>das Hochzeitsgeschenk</i>		<i>Geschenk</i>	GE-→*M,*F	✓	✓
<i>das Indossament</i>	-MENT→*M,*F SEM. EQUIV.→*M,*N (DIE ÜBERTRAGUNG)			X	✓
<i>das Informel</i>	-EL(L) ['el]→*M,*F			✓	✓
<i>das Inkrafttreten</i>	CONVERSION→*M,*F -EN→*F			✓	✓
<i>das J</i>	LETTERS→*M,*F V _{SHORT STOP C} →*F,*N CVC MONO→*F			X	✓
<i>das Juchtenleder</i>	-ER→*F,*N LEATHER TYPES→*M,*F	<i>Leder</i>	-ER→*F,*N	X	X
<i>das Jüngstenrecht</i>		<i>Recht</i>	R- MONO→*F -FT/-CHT→*M CVCC MONO→*F	✓	✓
<i>das Kilovolt</i>	SCI. UNITS→*M,*F	<i>Volt</i>	SCI. UNITS→*M,*F -[[]C MONO→*F CVCC MONO→*F	✓	✓
<i>das Kontaktformular</i>	-AR→*M,*F	<i>Formular</i>	-AR→*M,*F	✓	✓
<i>das Kopfzerbrechen</i>	CONVERSION→*M,*F -EN→*F			✓	✓
<i>das Langhaus</i>	DWELL→*M	<i>Haus</i>	VV _{DIPHTH} C MONO→*F CVC MONO→*F DWELL→*M	✓	✓

² Justified in Zubin & Köpcke (1986: 166)

<i>das Leben</i>	CONVERSION→*M,*F -EN→*F			✓	✓
<i>das Liniment</i>	-MENT→*M,*F			✓	✓
<i>das Lormalphabet</i>		<i>Alphabet</i>		X	X
<i>das Lunarium</i>	-IUM→*M,*F			✓	✓
<i>das Maßhalten</i>	CONVERSION→*M,*F -EN→*F			✓	✓
<i>das Meerschweinchen</i>	-CHEN→*M,*F DOM. ANIMALS→*F MAMMALS→*F	<i>Schweinchen</i>	-CHEN→*M,*F DOM. ANIMALS→*F MAMMALS→*F	✓	✓
<i>das Mikrofarad</i>	SCI. UNITS→*M,*F	<i>Farad</i>	SCI. UNITS→*M,*F	✓	✓
<i>das Miteinander</i>	CONVERSION→*M,*F -ER→*F,*N			X	✓
<i>das Monomer</i>	SCI. UNITS→*M,*F			✓	✓
<i>das Motiv</i>	-IV→*M,*F			✓	✓
<i>das Nomen proprium</i>	-EN→*F WORD CLASSES→*M,*F	<i>Nomen</i>	-EN→*F WORD CLASSES→*M,*F	✓	✓
<i>das Normalmaß</i>		<i>Maß</i>	V _{LONG} (C) MONO→*F CVC MONO→*F SUP.→*M,*F ³	✓	✓
<i>das Nuteisen</i>	-EN→*F STICK→*N	<i>Eisen</i>	METALS→*M,*F ELEMENTS→*M,*F -EN→*F	✓	✓
<i>das Oktavformat</i>	-AT→*M,*F	<i>Format</i>	-AT→*M,*F	✓	✓
<i>das Osmium</i>	METALS→*M,*F -IUM→*M,*F ELEMENTS→*M,*F			✓	✓
<i>das Partizip</i>	-IP [i:p]→*M,*F			✓	✓
<i>das Perpetuum mobile</i>	-UM→*M,*F	<i>Perpetuum</i>	-UM→*M,*F	✓	✓
<i>das Polarisationsmikroskop</i>	-SKOP→*M,*F	<i>Mikroskop</i>	-SKOP→*M,*F	✓	✓
<i>das Postgiroamt</i>		<i>Amt</i>	INSTITUTIONS→*M SUP→*M,*F	✓	✓
<i>das Q</i>	LETTERS→*M,*F V _{LONG} (C) MONO→*F CV MONO→*N			X	✓
<i>das Rauschgold</i>	CHEM. COMP./SUBST.→*M,*F	<i>Gold</i>	METALS→*M,*F ELEMENTS→*M,*F -[[]C MONO→*F CVCC MONO→*F	✓	✓
<i>das Rektapapier</i>	-IER /i:r/ POLY.→*F	<i>Papier</i>	-IER /i:r/ POLY.→*F	X	X
<i>das Rondeau</i>	-EAU→*M,*F			✓	✓
<i>das Rubidium</i>	METALS→*M,*F -IUM→*M,*F ELEMENTS→*M,*F			✓	✓
<i>das Säumchen</i>	-CHEN→*M,*F			✓	✓
<i>das Schaffen</i>	CONVERSION→*M,*F -EN→*F			✓	✓
<i>das Schembartlaufen</i>	CONVERSION→*M,*F -EN→*F			✓	✓
<i>das Schienbein</i>	STICK→*N	<i>Bein</i>	-N(C)(C)→*F,*N VV _{DIPHTH} C MONO→*F CVC MONO→*F	X	X
<i>das Schlittschuhlaufen</i>	CONVERSION→*M,*F -EN→*F GAMES→*M,*F			✓	✓
<i>das Schnällchen</i>	-CHEN→*M,*F			✓	✓

³ Justified in Zubin & Köpcke (1986: 166)

<i>das Schnippchen</i>	-CHEN→*M,*F			✓	✓
<i>das Sophisma</i>	ISMA→*M,*F -A→*M,*N			X	✓
<i>das Spektroskop</i>	-SKOP→*M,*F			✓	✓
<i>das Stop-and-go</i>	CONVERSION→*M,*F			✓	✓
<i>das Supremum</i>	-UM→*M,*F			✓	✓
<i>das Tamtam</i>	MUS. INSTR.→*M SEM. EQUIV.→*F (DER/DAS GONG)			✓	✓
<i>das Tempo</i>	-O→*F MUS. TERM/COMP→*M SEM. EQUIV.→*M,*N (DIE GESCHWINDIGKEIT)			X	✓
<i>das Tipi</i>	DWELL→*M SEM. EQUIV.→*M,*F (DAS INDIANERZELT)			✓	✓
<i>das Triolett</i>	-ETT→*M,*F			✓	✓
<i>das Uhrband</i>		<i>Band</i>	CVCC MONO→*F -N(C)(C)→*F,*N DEVERBAL STEM→*F,*N	X	X
<i>das Universalmittel</i>		<i>Mittel</i>	SUP.→*M,*F	✓	✓
<i>das UrhG = Urheberrechtsgesetz</i>	GE→*M,*F	<i>Gesetz</i>	GE→*M,*F	✓	✓
<i>das Vadium</i>	-IUM→*M,*F			✓	✓
<i>das Vorstadium</i>	-IUM→*M,*F	<i>Stadium</i>	-IUM→*M,*F	✓	✓
<i>das Waltharilied</i>	MUS. TERM/COMP→*M	<i>Lied</i>	V _{LONG} (C) MONO→*F CVC MONO→*F MUS. TERM/COMP→*M	✓	✓
<i>das Weißbuch</i>		<i>Buch</i>	CVC MONO→*F	X	X
<i>das Wissen</i>	CONVERSION→*M,*F -EN→*F SUP.→*M,*F ⁴			✓	✓
<i>das Wohlverhalten</i>	CONVERSION→*M,*F -EN→*F			✓	✓
<i>das ZDF = Zweites Deutsches Fernsehen</i>	CONVERSION→*M,*F -EN→*F	<i>Fernsehen</i>	CONVERSION→*M,*F -EN→*F	✓	✓
<i>das Zeitunglesen</i>	CONVERSION→*M,*F -EN→*F			✓	✓
<i>das Zusammenwirken</i>	CONVERSION→*M,*F -EN→*F			✓	✓
<i>das Zyan</i>	CHEM COMP/SUB→*M,*F -AN→*M,*F			✓	✓
<i>der Abend</i>	TIMES OF DAY→*F,*N			✓	✓
<i>der Achter</i>	-ER SUFF.→*F,*N			✓	✓
<i>der Affenbrotbaum</i>	TREES→*M,*N	<i>Baum</i>	-N(C)(C)→*F,*N VV _{DIPHTHC} MONO→*F CVC MONO→*F	✓	✓
<i>der Angelpunkt</i>		<i>Punkt</i>	-N(C)(C)→*F,*N CVCCC MONO→*F,*N	✓	✓
<i>der Anreiz</i>		<i>Anreiz</i>		✓	✓
<i>der Ara</i>	-A→*M,*N BIRDS→*F			✓	✓
<i>der Areopag</i>				✓	✓
<i>der Asphalt</i>	CHEM. COMP./SUBST.→*M,*F -ALT→*N			✓	✓

⁴ Justified in Zubin & Köpcke (1986: 166)

<i>der Audiostream</i>	SEM. EQUIV. → *F, *N (DER DATENSTROM)			✓	✓
<i>der Aufzugsschacht</i>		<i>Schacht</i>	-FT/-CHT → *M CVCC MONO → *F	X	X
<i>der Ausdrusch</i>	DEVERBAL STEM → *F, *N			✓	✓
<i>der Ausheber</i>	-ER SUFF. → *F, *N			✓	✓
<i>der Avistawechsel</i>		<i>Wechsel</i>	DEVERBAL STEM → *F, *N	✓	✓
<i>der Backenbart</i>		<i>Bart</i>	V _{LONG} (C) MONO → *F -/r/ STOP (C) MONO → *F, *N CVCC MONO → *F	✓	✓
<i>der Bautzener</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Bedrücker</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Beigeschmack</i>	GE → *M, *F	<i>Geschmack</i>	GE → *M, *F DEVERBAL STEM → *F, *N	✓	✓
<i>der Bergführer</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N	<i>Führer</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N	✓	✓
<i>der Besatz</i>	DEVERBAL STEM → *F, *N			✓	✓
<i>der Beutenhonig</i>		<i>Honig</i>	-IG/-ICH → *F, *N	✓	✓
<i>der Bezwinger</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Brandy</i>	ALC. DRINKS → *F, *N SEM. EQUIV. → *F, *N (DER WEINBRAND)			✓	✓
<i>der Bronzit</i>	(-IT → *F, *N / -IT → *F) ROCKS/MINERALS → *F, *N			✓	✓
<i>der Bummelant</i>	-ANT → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Bürgersteig</i>		<i>Steig</i>	[]C- MONO → *F, *N VV _{DIPHTH} C MONO → *F CCVV _{DIPHTHONG} - MONO → *F CCVC MONO → *F, *N DEVERBAL STEM → *F, *N	✓	✓
<i>der Byzantiner</i>	-INER → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Chickenwing</i>	SEM. EQUIV. → *F, *N (DER HÄHNCHENFLÜGEL)			✓	✓
<i>der Containerverkehr</i>		<i>Verkehr</i>	(DEVERBAL STEM → *F, *N)	✓	✓
<i>der DAAD = Deutscher Akademischer Austauschdienst</i>	INSTITUTIONS → *M	<i>Dienst</i>	CVCCC MONO → *F, *N -N(C)(C) → *F, *N D- MONO → *F V _{LONG} (C) MONO → *F -ST/-SCHT/-ZT → *N	✓	✓
<i>der Defroster</i>	-ER SUFF. → *F, *N SEM. EQUIV. → *F, *N (DER ENTEISER)			✓	✓
<i>der Dienstgebrauch</i>		<i>Gebrauch</i>	GE → *M, *F DEVERBAL STEM → *F, *N	✓	✓
<i>der Dime</i>	D- MONO → *F -N(C)(C) MONO → *F, *N VV _{DIPHTH} C MONO → *F CVC MONO → *F MON. UNITS → *F, *N SEM. EQUIV. → *F, *N (DER CENT)			✓	✓
<i>der Dipteros</i>				✓	✓

<i>der Docker</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N SEM. EQUIV. → *F, *N (DER HAFENARBEITER)			✓	✓
<i>der Dom</i>	D- MONO → *F -N(C)(C) MONO → *F, *N V _{LONG} (C) MONO . → *F CVC MONO → *F			✓	✓
<i>der Dorfclub, Dorfklub</i>	GROUPS → *M, *N -B → *F SEM. EQUIV. → *F, *N (DER DORFVEREIN)	<i>Club, Klub</i>	CCVC MONO → *F, *N -B → *F GROUPS → *M, *N SEM. EQUIV. → *F, *N (DER VEREIN)	✓	✓
<i>der Drall</i>	-L MONO → *F CCVC MONO → *F, *N DR- → *F, *N			✓	✓
<i>der Dreiklang</i>	MUS. TERM/COMP → *M -ANG → *F, *N	<i>Klang</i>	-N(C)(C) → *F, *N CCVC MONO → *F, *N DEVERBAL STEM → *F, *N -ANG → *F, *N	✓	✓
<i>der Dual</i>	-AL → *F			✓	✓
<i>der Dweil</i>	-L MONO → *F VV _{DIPHTHC} MONO → *F CCVV _{DIPHTHONG} - MONO → *F CCVC MONO → *F, *N			✓	✓
<i>der Ehestreit</i>		<i>Streit</i>	VV _{DIPHTHC} MONO → *F	✓	✓
<i>der Einriss</i>		<i>Riss</i>	CVC MONO → *F R- MONO → *F DEVERBAL STEM → *F, *N	✓	✓
<i>der Einstundentakt</i>	UNITS OF TIME → *N	<i>Takt</i>	T- MONO → *F V _{SHORT STOP C} → *F, *N CVCC MONO → *F	✓	✓
<i>der Eisenbahner</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Encoder, Enkoder</i>	-ER SUFF. → *F, *N SEM. EQUIV. → *F, *N (DER GEBER)			✓	✓
<i>der Eremit</i>	(-IT → *F, *N / -IT → *F) GEN. PERSON → *F, *N			✓	✓
<i>der Erweiterungsbau</i>		<i>Bau</i>	CV MONO → *N DEVERBAL STEM → *F, *N SUP. → *M, *F	✓	✓
<i>der Esslöffel</i>	STICK → *N	<i>Löffel</i>	STICK → *N	✓	✓
<i>der Eunuch</i>	MEN → *F, *N			✓	✓
<i>der Falschspieler</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N	<i>Spieler</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N	✓	✓
<i>der Federling</i>	-LING → *F, *N INSECTS → *M, *N			✓	✓
<i>der Feldverweis</i>		<i>Verweis</i>	DEVERBAL STEM → *F, *N	✓	✓
<i>der Fetischist</i>	-IST → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Fladerschnitt</i>		<i>Schnitt</i>	[[C]- MONO → *F, *N CCVC MONO → *F, *N DEVERBAL STEM → *F, *N	✓	✓
<i>der Fluchtversuch</i>		<i>Versuch</i>	DEVERBAL STEM → *F, *N	✓	✓
<i>der Folklorist</i>	-IST → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Fraktionssprecher</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N	<i>Sprecher</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N	✓	✓

<i>der Fremdstoff</i>		<i>Stoff</i>	[]C- MONO→*F,*N CCVC MONO→*F,*N SUP.→*M,*F	✓	✓
<i>der F-Schlüssel</i>	MUS. TERM/COMP→*M	<i>Schlüssel</i>	STICK→*N	✓	✓
<i>der Galvaniseur</i>	-EUR/ÖR→*F,*N/ -EUR/ÖR→*N GEN. PERSON→*F,*N			✓	✓
<i>der Geograf, Geograph</i>	GEN. PERSON→*F,*N			✓	✓
<i>der Geschäftsschluss</i>		<i>Schluss</i>	[]C- MONO→*F,*N CCVC MONO→*F,*N DEVERBAL STEM→*F,*N	✓	✓
<i>der Glockenmantel</i>		<i>Mantel</i>		✓	✓
<i>der Gottmensch</i>	GEN. PERSON→*F,*N	<i>Mensch</i>	-(c)(c)// MONO →*F,*N -N(c)(c)→*F,*N GEN. PERSON→*F,*N CVCC MONO→*F	✓	✓
<i>der Großbuchstabe</i>	-E→*M,*N	<i>Buchstabe</i>	-E→*M,*N SUP→*M,*F	x	x
<i>der Grufti</i>	GEN. PERSON→*F,*N			✓	✓
<i>der Handball</i>	GAMES→*M,*F	<i>Ball</i>	-L MONO→*F CVC MONO→*F	✓	✓
<i>der Hannoveraner</i>	-ANER→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Hasenpfeffer</i>	-ER→*F,*N	<i>Pfeffer</i>	-ER→*F,*N	✓	✓
<i>der Hausmeister</i>	-ER→*F,*N GEN. PERSON→*F,*N	<i>Meister</i>	-ER→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Heißhunger</i>		<i>Hunger</i>	-ER→*F,*N	✓	✓
<i>der Helvetier</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Herumtreiber</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Treiber</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Himmelskörper</i>	-ER→*F,*N SUP.→*M,*F	<i>Körper</i>	-ER→*F,*N	✓	✓
<i>der Hochgenuss</i>		<i>Genuss</i>	GE→*M,*F DEVERBAL STEM→*F,*N	✓	✓
<i>der Hops</i>	-C S MONO →*F,*N V _{SHORT STOP C} →*F,*N CVCC MONO→*F DEVERBAL STEM→*F,*N			✓	✓
<i>der Hunderter</i>	-ER SUFF.→*F,*N			✓	✓
<i>der Hypothekenzins</i>		<i>Zins</i>	-C S MONO →*F,*N CCVCC→*F,*N -N(c)(c)→*F,*N	✓	✓
<i>der Intendant</i>	-ANT→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Interrogativsatz</i>		<i>Satz</i>	-C S MONO →*F,*N V _{SHORT STOP C} →*F,*N CVCC MONO→*F	✓	✓
<i>der Islam</i>	RELIGIONS→*F,*N (V _{STRESSED [m]} →*M,*F)			✓	✓
<i>der Jetztmensch</i>	GEN. PERSON→*F,*N	<i>Mensch</i>	-(c)(c)// MONO →*F,*N -N(c)(c)→*F,*N GEN. PERSON→*F,*N CVCC MONO→*F	✓	✓
<i>der Kalamit</i>	(-IT→*F,*N/-IT→*F) PLANTS→*N			✓	✓
<i>der Kapuzineraffe</i>	-E→*M,*N MAMMALS→*F PRIMATES→*F,*N	<i>Affe</i>	-E→*M,*N MAMMALS→*F PRIMATES→*F,*N	✓	✓

<i>der Karlsbader</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Kauerstart</i>		<i>Start</i>	[]C- MONO → *F, *N CCVCC → *F, *N -/r/ STOP (C) MONO → *F, *N DEVERBAL STEM → *F, *N	✓	✓
<i>der Key-Account-Manager</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N SEM. EQUIV. → *F, *N (DER CHEF)	<i>Manager</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N SEM. EQUIV. → *F, *N (DER CHEF)	✓	✓
<i>der Kirschgeist</i>	ALC. DRINKS → *F, *N	<i>Geist</i>	-ST/-SCHT/-ZT → *N CVCC MONO → *F	✓	✓
<i>der Knacker</i>	-ER SUFF. → *F, *N MEN → *F, *N			✓	✓
<i>der Kohlenherd</i>		<i>Herd</i>	V _{LONG} (C) MONO → *F -/r/ STOP (C) MONO → *F, *N CVCC MONO → *F	✓	✓
<i>der Konformismus</i>	-ISMUS → *F, *N			✓	✓
<i>der Koton</i>	FABRIC → *F, *N SEM. EQUIV. → *M, *N (DIE BAUMWOLLE) / → *F (DER/DAS COTTON)			✓	✓
<i>der Kreisdurchmesser</i>	-ER SUFF. → *F, *N	<i>Messer</i>	-ER SUFF. → *F, *N	✓	✓
<i>der Krumme</i>	-E → *M, *N MAMMALS → *F			✓	✓
<i>der Kühlwagen</i>	-EN → *F	<i>Wagen</i>	-EN → *F	✓	✓
<i>der Landadel</i>		<i>Adel</i>		✓	✓
<i>der Läufer</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Leichenwagen</i>	-EN → *F	<i>Wagen</i>	-EN → *F	✓	✓
<i>der Literaturnobelpreis</i>		<i>Preis</i>	VV _{DIPHTH} C MONO → *F CCVV _{DIPHTHONG} - MONO → *F CCVC MONO → *F, *N	✓	✓
<i>der Löffelstiel</i>	STICK → *N	<i>Stiel</i>	[]C- MONO → *F, *N -L MONO → *F V _{LONG} (C) MONO → *F CCVC MONO → *F, *N STICK → *N	✓	✓
<i>der Lorenzstrom</i>	RIVERS NON-GER. → *F, *N RIVERS NON-EUROPE → *F, *N	<i>Strom</i>	-N(C)(C) → *F, *N V _{LONG} (C) MONO → *F CCCVC MONO → *F, *N	✓	✓
<i>der Löwenbändiger</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N	<i>Bändiger</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N	✓	✓
<i>der Mainzer</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Mandelbaum</i>	TREES → *M, *N	<i>Baum</i>	-N(C)(C) → *F, *N VV _{DIPHTH} C MONO → *F CVC MONO → *F	✓	✓
<i>der Manteltarif</i>		<i>Tarif</i>		✓	✓
<i>der Martinstag</i>	DAYS → *F, *N	<i>Tag</i>	T- MONO → *F V _{LONG} (C) MONO → *F CVC MONO → *F UNITS OF TIME → *N	✓	✓
<i>der Matsch-und-Schnee-Reifen</i>	-EN → *F	<i>Reifen</i>	-EN → *F	✓	✓
<i>der Meisterdetektiv</i>	-IV → *M, *F GEN. PERSON → *F, *N	<i>Detektiv</i>	-IV → *M, *F GEN. PERSON → *F, *N	✓	✓
<i>der Menschengest</i>		<i>Geist</i>	-ST/-SCHT/-ZT → *N CVCC MONO → *F	✓	✓

<i>der Metropolit</i>	(-IT→*F,*N/-IT→*F) GEN. PERSON→*F,*N			✓	✓
<i>der Mischer</i>	-ER SUFF.→*F,*N			✓	✓
<i>der Mittelstürmer</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Stürmer</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Molekularbiologe</i>	-(O)LOGE→*F,*N GEN. PERSON→*F,*N	<i>Biologe</i>	-(O)LOGE→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Moralprediger</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Prediger</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Mustang</i>	-ANG→*F,*N MAMMALS→*F			✓	✓
<i>der Nachschlag</i>	DEVERBAL STEM→*F,*N	<i>Schlag</i>	[]C-MONO→*F,*N V _{LONG} (C) MONO→*F CCVC MONO→*F,*N DEVERBAL STEM→*F,*N	✓	✓
<i>der Nagel</i>	STICK→*N			✓	✓
<i>der Nanga Parbat</i>	MOUNTAINS→*F,*N			✓	✓
<i>der Nationalstaat</i>		<i>Staat</i>	[]C-MONO→*F,*N V _{LONG} (C) MONO→*F CCVC MONO→*F,*N SUP.→*M,*F ⁵	✓	✓
<i>der Nebenbuhler</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Buhler</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der New-Orleans-Jazz</i>	MUSIC TYPES→*F,*N	<i>Jazz</i>	MUSIC TYPES→*F,*N (-C S MONO→*F,*N) (V _{SHORT STOP C} →*F,*N) (CVCC MONO→*F) (CCVC MONO→*F,*N)	✓	✓
<i>der Offiziersanwärter</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Anwärter</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Out-of-area-Einsatz</i>		<i>Einsatz</i>	DEVERBAL STEM→*F,*N	✓	✓
<i>der Patagonier</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Petent</i>	-ENT→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Pfifferling</i>	-LING→*F,*N			✓	✓
<i>der Phönizier</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Pickhammer</i>	-ER→*F,*N	<i>Hammer</i>	-ER→*F,*N	✓	✓
<i>der Platzanweiser</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Anweiser</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Porphyr</i>	ROCKS/MINERALS→*F,*N -ÜR/-ÜHR /y:r/→*M,*N			✓	✓
<i>der Präfekt</i>	-EKT→*F GEN. PERSON→*F,*N			✓	✓
<i>der Presenter</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N SEM. EQUIV.→*F,*N (DER SPRECHER)			✓	✓
<i>der Prolaps</i>				✓	✓
<i>der Publikumsgeschmack</i>	GE→*M,*F	<i>Geschmack</i>	GE→*M,*F DEVERBAL STEM→*F,*N	✓	✓
<i>der Rähm</i>	-N(C)(C) MONO→*F,*N R- MONO→*F V _{LONG} (C) MONO→*F CVC MONO→*F			✓	✓
<i>der Rapfen</i>	-EN→*F FISH→*F,*N			✓	✓

⁵ Justified in Zubin & Köpcke (1986: 166)

<i>der Rechercheauftrag</i>		<i>Auftrag</i>		✓	✓
<i>der Reinfeld</i>		<i>Fall</i>	-L MONO→*F (DEVERBAL STEM→*F,*N) CVC MONO→*F	✓	✓
<i>der Rotang</i>	TREES→*M,*N -ANG→*F,*N			✓	✓
<i>der Rufer</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Samländer</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Satellit</i>	(-IT→*F,*N/-IT→*F)			✓	✓
<i>der Schanker</i>	-ER→*F,*N (DISEASES→*M/DISEASES →*N)			✓	✓
<i>der Schiri = Schiedsrichter</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Richter</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Schlangenbeschwörer</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Beschwörer</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Schlemmer</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Schmätzer</i>	-ER SUFF.→*F,*N BIRDS→*F			✓	✓
<i>der Schönfärber</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Färber</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Schrankkoffer</i>		<i>Koffer</i>	-ER→*F,*N	✓	✓
<i>der Schrumpfergermane</i>	-E→*M,*N GEN. PERSON→*F,*N	<i>Germane</i>	-E→*M,*N GEN. PERSON→*F,*N	✓	✓
<i>der Schund</i>	-N(C)(C) MONO→*F,*N CVCC MONO→*F WASTE→*F,*N			✓	✓
<i>der Sechsling</i>	-LING→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Sekans</i>				✓	✓
<i>der Siderit</i>	(-IT→*F,*N/-IT→*F) ROCKS/MINERALS→*F,*N			✓	✓
<i>der Sikahirsch</i>	DOM. ANIMALS→*F MAMMALS→*F	<i>Hirsch</i>	-(C)(C)// MONO→*F,*N CVCC MONO→*F DOM. ANIMALS→*F MAMMALS→*F	✓	✓
<i>der Sinnenrausch</i>		<i>Rausch</i>	-(C)(C)// MONO→*F,*N R- MONO→*F VV _{DIPHTH} C MONO→*F CVC MONO→*F	✓	✓
<i>der Skatabend</i>		<i>Abend</i>	TIMES OF DAY→*F,*N	✓	✓
<i>der Slowake</i>	-E→*M,*N GEN. PERSON→*F,*N			✓	✓
<i>der Sog</i>	V _{LONG} (C) MONO→*F CVC MONO→*F			✓	✓
<i>der Sommer</i>	SEASONS→*F,*N -ER→*F,*N			✓	✓
<i>der Sparerfreibetrag</i>		<i>Betrag</i>	DEVERBAL STEM→*F,*N	✓	✓
<i>der Spitzkühler</i>	-ER SUFF.→*F,*N	<i>Kühler</i>	-ER SUFF.→*F,*N	✓	✓
<i>der Standartenträger</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Träger</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Statist</i>	-IST→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Stelzvogel</i>	BIRDS→*F	<i>Vogel</i>		✓	✓
<i>der Steuerfreibetrag</i>		<i>Betrag</i>	DEVERBAL STEM→*F,*N	✓	✓
<i>der Strampelanzug</i>		<i>Anzug</i>	DEVERBAL STEM→*F,*N	✓	✓

<i>der Suizident</i>	-ENT→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Tausender</i>	-ER SUFF.→*F,*N			✓	✓
<i>der Test</i>	-ST/-SCHT/-ZT→*N T- MONO→*F CVCC MONO→*F SEM. EQUIV.→*M,*N (DIE UNTERSUCHUNG)			✓	✓
<i>der Thyristor</i>	-OR→*F,*N			✓	✓
<i>der Topverdiener</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Verdiener</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Tourenski, Tourenski</i>		<i>Ski, Schi</i>	CV MONO→*N V _{LONG} (C) MONO→*F	✓	✓
<i>der Tran</i>	-N(C)(C) MONO→*F,*N V _{LONG} (C) MONO→*F CCVC MONO→*F,*N			✓	✓
<i>der Umformer</i>	-ER SUFF.→*F,*N			✓	✓
<i>der Undank</i>		<i>Dank</i>	D- MONO→*F -N(C)(C)→*F,*N CVCC MONO→*F	✓	✓
<i>der Ungeist</i>		<i>Geist</i>	-ST/-SCHT/-ZT→*N CVCC MONO→*F	✓	✓
<i>der Unsegen</i>	-EN→*F	<i>Segen</i>	-EN→*F	✓	✓
<i>der Unterführer</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Führer</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Vegetarier</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Verbesserer</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Verdross</i>	DEVERBAL STEM→*F,*N			✓	✓
<i>der Verkauf</i>	DEVERBAL STEM→*F,*N	<i>Kauf</i>	VV _{DIPHTH} C MONO→*F CVC MONO→*F DEVERBAL STEM→*F,*N	✓	✓
<i>der Vielflieger</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	<i>Flieger</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N	✓	✓
<i>der Vogelschutz</i>		<i>Schutz</i>	-C S MONO→*F,*N V _{SHORT} STOP C→*F,*N CVCC MONO→*F DEVERBAL STEM→*F,*N	✓	✓
<i>der Vollender</i>	-ER SUFF.→*F,*N GEN. PERSON→*F,*N			✓	✓
<i>der Voraus</i>	CONVERSION→*M,*F			X	X
<i>der Wartestand</i>		<i>Stand</i>	[]C- MONO→*F,*N CCVCC→*F,*N -N(C)(C)→*F,*N DEVERBAL STEM→*F,*N	✓	✓
<i>der Webstuhl</i>		<i>Stuhl</i>	[]C- MONO→*F,*N -L MONO→*F V _{LONG} (C) MONO→*F CCVC MONO→*F,*N	✓	✓
<i>der Wehrpass</i>		<i>Pass</i>	CVC MONO→*F	✓	✓
<i>der Weitschuss</i>		<i>Schuss</i>	CVC MONO→*F DEVERBAL STEM→*F,*N	✓	✓
<i>der Werkbus</i>		<i>Bus</i>	CVC MONO→*F	✓	✓
<i>der Wettbewerbsnachteil</i>		<i>Teil</i>	T- MONO→*F VV _{DIPHTH} C MONO→*F -L MONO→*F CVC MONO→*F	✓	✓

<i>der Wiedehopf</i>	BIRDS→*F			✓	✓
<i>der Windhund</i>	DOM. ANIMALS→*F MAMMALS→*F	<i>Hund</i>	-N(C)(C)→*F,*N CVCC MONO→*F DOM. ANIMALS→*F MAMMALS→*F	✓	✓
<i>der Zirkelkasten</i>	-EN→*F	<i>Kasten</i>	-EN→*F	✓	✓
<i>der Zusammenbruch</i>	DEVERBAL STEM→*F,*N	<i>Bruch</i>	CCVC MONO→*F,*N DEVERBAL STEM→*F,*N	✓	✓
<i>der Zweiteiler</i>	-ER SUFF.→*F,*N			✓	✓
<i>die ABM = Arbeitsbeschaffungsmaßna hme</i>	-E→*M,*N	<i>Maßnahme</i>	-E→*M,*N	✓	✓
<i>die Abseitsposition</i>	-ION→*M,*N	<i>Position</i>	-ION→*M,*N	✓	✓
<i>die Abwärme</i>	-E→*M,*N	<i>Wärme</i>	-E→*M,*N	✓	✓
<i>die Adduktion</i>	-ION→*M,*N			✓	✓
<i>die Agrogentechnik, Agro- Gentechnik</i>	-IK→*M,*N	<i>Technik</i>	-IK→*M,*N	✓	✓
<i>die Akrozephalie</i>	-IE ['i:]→*M,*N (DISEASES→*M/DISEASES →*N)			✓	✓
<i>die Alchimie</i>	-IE ['i:]→*M,*N			✓	✓
<i>die All-Star-Band</i>	GROUPS→*M,*N	<i>Band</i>	GROUPS→*M,*N CVCC MONO→*F -N(C)(C)→*F,*N SEM. EQUIV.→*M,*N (DIE GRUPPE)	X	✓
<i>die Alternanz</i>	-ANZ→*M,*N			✓	✓
<i>die Amelioration</i>	-ION→*M,*N			✓	✓
<i>die Anknüpfung</i>	-UNG→*M,*N			✓	✓
<i>die Antwort</i>				X	X
<i>die Apokope</i>	-E→*M,*N			✓	✓
<i>die Aszetik</i>	-IK→*M,*N			✓	✓
<i>die Aufgliederung</i>	-UNG→*M,*N	<i>Gliederung</i>	-UNG→*M,*N	✓	✓
<i>die Aufsichtung</i>	-UNG→*M,*N			✓	✓
<i>die Ausreifung</i>	-UNG→*M,*N			✓	✓
<i>die Äußerung</i>	-UNG→*M,*N			✓	✓
<i>die Autarkie</i>	-IE ['i:]→*M,*N			✓	✓
<i>die Bandage</i>	-AGE→*M,*N			✓	✓
<i>die Bärentraube</i>	-E→*M,*N FLOWERS→*M,*N PLANTS→*N	<i>Traube</i>	FRUITS→*M,*N -E→*M,*N	✓	✓
<i>die Bassgeige</i>	-E→*M,*N MUS. INSTR.→*M	<i>Geige</i>	-E→*M,*N MUS. INSTR.→*M	✓	✓
<i>die Begriffsbestimmung</i>	-UNG→*M,*N	<i>Bestimmung</i>	-UNG→*M,*N	✓	✓
<i>die Bekohlung</i>	-UNG→*M,*N			✓	✓
<i>die Bessemerbirne</i>	-E→*M,*N	<i>Birne</i>	-E→*M,*N FRUITS→*M,*N	✓	✓
<i>die Bildhauerkunst</i>		<i>Kunst</i>	ACAD. DISCIPLINES→*M,*N -ST/-SCHT/-ZT→*N -N(C)(C)→*F,*N CVCCC MONO→*F,*N	X	✓
<i>die Biomedizin</i>	-IN [i:n]→*M,*F ACAD. DISCIPLINES→*M,*N	<i>Medizin</i>	-IN [i:n]→*M,*F ACAD. DISCIPLINES→*M,*N	✓	✓
<i>die Blausäure</i>	-E→*M,*N CHEM. COMP./SUBST.→*M,*F	<i>Säure</i>	-E→*M,*N CHEM. COMP./SUBST.→*M,*F	✓	✓
<i>die Bloßstellung</i>	-UNG→*M,*N	<i>Stellung</i>	-UNG→*M,*N	✓	✓

<i>die Bodybuilderin</i>	-IN SUFF. → *M, *N WOMEN → *M, *N SEM. EQUIV. → *M, *N (DIE MUSKELFRAU)			✓	✓
<i>die Bougainvillea</i>	-A → *M, *N FLOWERS → *M, *N TREES → *M, *N PLANTS → *N			✓	✓
<i>die Bremsrakete</i>	-E → *M, *N	<i>Rakete</i>	-E → *M, *N	✓	✓
<i>die Buchführung</i>	-UNG → *M, *N	<i>Führung</i>	-UNG → *M, *N	✓	✓
<i>die Cantate, Kantate</i>	-E → *M, *N MUS. TERM/COMP → *M			✓	✓
<i>die C-Dur-Tonleiter</i>	-ER → *F, *N MUS. TERMS → *M	<i>Leiter</i>	-ER → *F, *N	x	x
<i>die Chargennummer</i>	-ER → *F, *N	<i>Nummer</i>	-ER → *F, *N	x	x
<i>die Chromatografie, -phie</i>	-IE ['i:] → *M, *N			✓	✓
<i>die Cranberry</i>	FRUITS → *M, *N SEM. EQUIV. → *M, *N (DIE PREISELBEERE)			✓	✓
<i>die Dahlie</i>	-IE [iə] → *M, *N FLOWERS → *M, *N PLANTS → *N			✓	✓
<i>die Daphnia</i>	-A → *M, *N			✓	✓
<i>die Dasselbeule</i>	-E → *M, *N (DISEASES → *M/DISEASES → *N)	<i>Beule</i>	-E → *M, *N	✓	✓
<i>die Dendrochronologie</i>	-IE ['i:] → *M, *N	<i>Chronologie</i>	-IE ['i:] → *M, *N	✓	✓
<i>die Disproportion</i>	-ION → *M, *N	<i>Proportion</i>	-ION → *M, *N	✓	✓
<i>die Dunstglocke</i>	-E → *M, *N	<i>Glocke</i>	-E → *M, *N MUS. INSTR. → *M	✓	✓
<i>die Durchhalteparole</i>	-E → *M, *N	<i>Parole</i>	-E → *M, *N	✓	✓
<i>die Durchschrift</i>		<i>Schrift</i>	-FT/-CHT → *M CCVCC → *F, *N [[C-MONO → *F, *N	x	✓
<i>die Einehe</i>	-E → *M, *N	<i>Ehe</i>	-E → *M, *N	✓	✓
<i>die Einigkeit</i>	-HEIT/-(IG)KEIT → *M, *N			✓	✓
<i>die Elektrizität</i>	-TÄT → *M, *N			✓	✓
<i>die EM = Europameisterschaft</i>	-SCHAFT → *M, *N	<i>Meisterschaft</i>	-SCHAFT → *M, *N	✓	✓
<i>die Enkelin</i>	-IN SUFF. → *M, *N WOMEN → *M, *N			✓	✓
<i>die Entkriminalisierung</i>	-UNG → *M, *N	<i>Entkriminalisierung</i>	-UNG → *M, *N	✓	✓
<i>die Entvölkerung</i>	-UNG → *M, *N			✓	✓
<i>die Erkundigung</i>	-UNG → *M, *N			✓	✓
<i>die Examensangst</i>		<i>Angst</i>	VCCC MONO → *M -ST/-SCHT/-ZT → *N	✓	✓
<i>die Expertengruppe</i>	-E → *M, *N GROUPS *M, *N	<i>Gruppe</i>	-E → *M, *N	✓	✓
<i>die Exzellenzinitiative</i>	-IVE → *M, *N	<i>Initiative</i>	-IVE → *M, *N	✓	✓
<i>die Finsternis</i>	-NIS → *M			✓	✓
<i>die Fortifikation</i>	-ION → *M, *N			✓	✓
<i>die Frittenbude</i>	-E → *M, *N	<i>Bude</i>	-E → *M, *N DWELL → *M	✓	✓
<i>die Fustage</i>	-AGE → *M, *N			✓	✓
<i>die Garonne</i>	-NNE [N(ə)] → *M, *N RIVERS NON-GER. → *F, *N RIVERS EUROPE → *N			x	✓

die Gedunsenheit	-HEIT/-(IG)KEIT→*M,*N GE→*M,*F			✓	✓
die Geisel	GEN. PERSON→*F,*N			x	x
die Generalagentin	-IN SUFF.→*M,*N WOMEN→*M,*N	Agentin	-IN SUFF.→*M,*N WOMEN→*M,*N	✓	✓
die Gerichtsmedizin	-IN [i:n]→*M,*F ACAD. DISCIPLINES→*M,*N	Medizin	-IN [i:n]→*M,*F ACAD. DISCIPLINES→*M,*N	✓	✓
die Gesetzgebung	-UNG→*M,*N			✓	✓
die Gewinnquote	-E→*M,*N	Quote	-E→*M,*N	✓	✓
die Gimpe	-E→*M,*N			✓	✓
die Glazialfauna	-A→*M,*N	Fauna	-A→*M,*N SUP.→*M,*F	✓	✓
die Grandel				x	x
die Gräzität	-TÄT→*M,*N			✓	✓
die GST = Gesellschaft für Sport und Technik	-SCHAFT→*M,*N GE→*M,*F GROUPS→*M,*N	Gesellschaft	-SCHAFT→*M,*N GE→*M,*F GROUPS→*M,*N	✓	✓
die Gutgläubigkeit	-HEIT/-(IG)KEIT→*M,*N			✓	✓
die Hairstylisten	-IN SUFF.→*M,*N WOMEN→*M,*N SEM. EQUIV.→*M,*N (DIE KÜNSTLERIN)			✓	✓
die Hegemonie	-IE ['i:]→*M,*N			✓	✓
die Herbst-Tagundnachtgleiche	-E→*M,*N	Gleiche	-E→*M,*N	✓	✓
die Hintereinanderschaltung	-UNG→*M,*N			✓	✓
die Holländerin	-IN SUFF.→*M,*N WOMEN→*M,*N			✓	✓
die Hubbrücke	-E→*M,*N	Brücke	-E→*M,*N	✓	✓
die Hybris	-IS→*M,*N			✓	✓
die Idiotin	-IN SUFF.→*M,*N WOMEN→*M,*N			✓	✓
die Imbezilität	-TÄT→*M,*N			✓	✓
die Inadäquatheit	-HEIT/-(IG)KEIT→*M,*N			✓	✓
die Insertion	-ION→*M,*N			✓	✓
die Inzahlungnahme	-E→*M,*N	Nahme	-E→*M,*N	✓	✓
die Janitscharenmusik	-IK→*M,*N MUSIC TYPES→*F,*N	Musik	-IK→*M,*N ACAD. DISCIPLINES→*M,*N	✓	✓
die Kabinettskrise	-E→*M,*N	Krise	-E→*M,*N	✓	✓
die Kaltblütigkeit	-HEIT/-(IG)KEIT→*M,*N			✓	✓
die Kampfmaßnahme	-E→*M,*N	Maßnahme	-E→*M,*N	✓	✓
die Kantonalbank	INSTITUTIONS→*M	Bank	CVCC MONO→*F -N(C)(C)→*F,*N INSTITUTIONS→*M	x	x
die Karwoche	-E→*M,*N	Woche	-E→*M,*N UNITS OF TIME→*N	✓	✓
die Kasusendung	-UNG→*M,*N	Endung	-UNG→*M,*N	✓	✓
die Kelle	-E→*M,*N STICK→*N			✓	✓
die Kimm	CVC MONO→*F -N(C)(C) MONO→*F,*N			x	x
die Klasse	-E→*M,*N			✓	✓
die Kleinschreibung	-UNG→*M,*N	Schreibung	-UNG→*M,*N	✓	✓
die Klinge	-E→*M,*N			✓	✓
die Knopflochseide	FABRIC→*F,*N -E→*M,*N	Seide	FABRIC→*F,*N -E→*M,*N	x	✓

die Kolmation	-ION→*M,*N			✓	✓
die Kommentatorin	-IN SUFF.→*M,*N WOMEN→*M,*N			✓	✓
die Komplexbrigade	-ADE→*M,*N GROUPS→*M,*N	Brigade	-ADE→*M,*N GROUPS→*M,*N	✓	✓
die Konnivenz	-ENZ→*M,*N			✓	✓
die Konvergenz	-ENZ→*M,*N			✓	✓
die Korpuskulartheorie	-IE ['i:]→*M,*N	Theorie	-IE ['i:]→*M,*N	✓	✓
die Krankenanstalt	INSTITUTIONS→*M -ALT→*N	Anstalt	INSTITUTIONS→*M SUP.→*M,*F ⁶ -ALT→*N	✓	✓
die Kriegsverbrecherin	-IN SUFF.→*M,*N WOMEN→*M,*N	Verbrecherin	-IN SUFF.→*M,*N WOMEN→*M,*N	✓	✓
die Kunst	ACAD. DISCIPLINES→*M,*N -ST/-SCHT/-ZT→*N -N(C)(C) MONO→*F,*N CVCCC MONO→*F,*N			x	✓
die Kursbewegung	-UNG→*M,*N	Bewegung	-UNG→*M,*N	✓	✓
die Lafette	-ETTE→*M,*N			✓	✓
die Laryngitis	-ITIS→*M,*N (DISEASES→*M/DISEASES →*N)			✓	✓
die Legalität	-TÄT→*M,*N			✓	✓
die Leiste	-E→*M,*N			✓	✓
die Lese	-E→*M,*N			✓	✓
die Liasformation	-ION→*M,*N	Formation	-ION→*M,*N	✓	✓
die Liederlichkeit	-HEIT/-(IG)KEIT→*M,*N			✓	✓
die Lyssa	-A→*M,*N (DISEASES→*M/DISEASES →*N)			✓	✓
die Magenschleimhaut		Haut	CVC MONO→*F VV _{DIPHTH} C MONO→*F (ORGANS→*M/ORGANS→ *N)	x	x
die Malaria	-A→*M,*N (DISEASES→*M/DISEASES →*N)			✓	✓
die Marke	-E→*M,*N			✓	✓
die MAZ = magnetische Bildaufzeichnung	-UNG→*M,*N	Aufzeichnung	-UNG→*M,*N	✓	✓
die Messerspitze	-E→*M,*N	Spitze	-E→*M,*N	✓	✓
die Mimin	-IN SUFF.→*M,*N WOMEN→*M,*N			✓	✓
die Modallogik	-IK→*M,*N	Logik	-IK→*M,*N	✓	✓
die Mühsal	-SAL→*M			✓	✓
die Munterkeit	-HEIT/-(IG)KEIT→*M,*N			✓	✓
die Nachäfferei	-(ER)EI→*M,*N	Äfferei	-(ER)EI→*M,*N	✓	✓
die Neotenie	-IE ['i:]→*M,*N			✓	✓
die Neubesetzung	-UNG→*M,*N	Besetzung	-UNG→*M,*N	✓	✓
die Niederdruckheizung	-UNG→*M,*N	Heizung	-UNG→*M,*N	✓	✓
die Nikotinhaltigkeit	-HEIT/-(IG)KEIT→*M,*N			✓	✓
die Noxe	-E→*M,*N			✓	✓
die Oberin	-IN SUFF.→*M,*N WOMEN→*M,*N			✓	✓
die Okkasion	-ION→*M,*N			✓	✓

⁶ Justified in Zubin & Köpcke (1986: 166)

die ÖNB = Österreichische Nationalbibliothek	-THEK→*M,*N	Bibliothek	-THEK→*M,*N	✓	✓
die Optionsanleihe	-E→*M,*N	Leihe	-E→*M,*N	✓	✓
die Orgelpfeife	-E→*M,*N STICK→*N	Pfeife	-E→*M,*N STICK→*N	✓	✓
die Pädagogisierung	-UNG→*M,*N			✓	✓
die Palpe	-E→*M,*N			✓	✓
die Panzerdivision	-ION→*M,*N GROUPS→*M,*N	Division	-ION→*M,*N GROUPS→*M,*N	✓	✓
die Paralalie	-IE ['i:]→*M,*N (DISEASES→*M/DISEASES →*N)			✓	✓
die Parkkralle	-E→*M,*N	Kralle	-E→*M,*N	✓	✓
die Pauschalität	-TÄT→*M,*N			✓	✓
die Pelzware	-E→*M,*N	Ware	-E→*M,*N SUP.→*M,*F	✓	✓
die Perfidie	-IE ['i:]→*M,*N			✓	✓
die Phänomenologie	-IE ['i:]→*M,*N			✓	✓
die Pimpernuss	FLOWERS→*M,*N PLANTS→*N	Nuss	CVC MONO→*F FRUITS→*M,*N	x	✓
die Plagiatorin	-IN SUFF.→*M,*N WOMEN→*M,*N			✓	✓
die Pluralisierung	-UNG→*M,*N			✓	✓
die Polymerisierung	-UNG→*M,*N			✓	✓
die Präsentation	-UNG→*M,*N			✓	✓
die Priorin	-IN SUFF.→*M,*N WOMEN→*M,*N			✓	✓
die Produktlinie	-IE [iə]→*M,*N	Linie	-IE [iə]→*M,*N	✓	✓
die Prospektierung	-UNG→*M,*N			✓	✓
die Prozessordnung	-UNG→*M,*N	Ordnung	-UNG→*M,*N	✓	✓
die Pupille	-ILLE→*M,*N			✓	✓
die Quarte	-E→*M,*N MUS. TERM/COMP→*M			✓	✓
die Quetschfalte		Falte	-E→*M,*N	✓	✓
die Rackerei	-(ER)EI→*M,*N			✓	✓
die Ratlosigkeit	-HEIT/-(IG)KEIT→*M,*N			✓	✓
die Rechtswissenschaft	-SCHAFT→*M,*N	Wissenschaft	-SCHAFT→*M,*N	✓	✓
die Regel				x	x
die Regung	-UNG→*M,*N			✓	✓
die Renitenz	-ENZ→*M,*N			✓	✓
die Reseda	-A→*M,*N FLOWERS→*M,*N PLANTS→*N			✓	✓
die Retourkutsche	-E→*M,*N	Kutsche	-E→*M,*N	✓	✓
die Rezeptur	-UR SUFF.→*M,*N			✓	✓
die Richtung	-UNG→*M,*N			✓	✓
die Rinne	-E→*M,*N			✓	✓
die Rockoper	-ER→*F,*N	Oper	MUSIC TYPES→*F,*N MUS. TERM/COMP→*M -ER→*F,*N	x	x
die Rupiah	-A→*M,*N (/IE [iə]→*M,*N) MON. UNITS→*F,*N			x	✓
die Sachlichkeit	-HEIT/-(IG)KEIT→*M,*N			✓	✓
die Salatkartoffel	VEG.ROOT→*M,*N	Kartoffel	VEG.ROOT→*M,*N	✓	✓

die Sanitärtechnik	-IK→*M,*N	Technik	-IK→*M,*N	✓	✓
die Schaufensterpuppe	-E→*M,*N	Puppe	-E→*M,*N	✓	✓
die Schwachstelle	-E→*M,*N	Stelle	-E→*M,*N	✓	✓
die Schwarzarbeit		Arbeit		x	x
die Schweißperle	-E→*M,*N	Perle	-E→*M,*N	✓	✓
die Schwestertracht		Tracht	-FT/-CHT→*M CCVCC→*F,*N	x	✓
die Sehnenzerrung	-UNG→*M,*N (DISEASES→*M/DISEASES →*N)	Zerrung	-UNG→*M,*N	✓	✓
die Selbstständigkeit	-HEIT/-(IG)KEIT→*M,*N			✓	✓
die Seniorengymnastik	GAMES→*M,*F -IK→*M,*N	Gymnastik	GAMES→*M,*F -IK→*M,*N	✓	✓
die Serologie	-IE ['i:]→*M,*N			✓	✓
die Shapingmaschine	-E→*M,*N	Maschine	-E→*M,*N	✓	✓
die Sowjetrepublik	-IK→*M,*N COUNTRIES→*M,*F	Republik	-IK→*M,*N	✓	✓
die Spielpause	-E→*M,*N	Pause	-E→*M,*N	✓	✓
die Sprachgrenze	-E→*M,*N	Grenze	-E→*M,*N	✓	✓
die Spundbohle	-E→*M,*N	Bohle	STICK→*N -E→*M,*N	✓	✓
die Stadtentwicklung	-UNG→*M,*N	Entwicklung	-UNG→*M,*N	✓	✓
die Steelband	GROUPS→*M,*N	Band	GROUPS→*M,*N CVCC MONO→*F -N(C)(C)→*F,*N SEM. EQUIV.→*M,*N (DIE GRUPPE)	x	✓
die Stilllegung, Still-Legung	-UNG→*M,*N			✓	✓
die Strenge	-E→*M,*N			✓	✓
die Strukturtapete	-E→*M,*N	Tapete	-E→*M,*N	✓	✓
die Stundenzahl		Zahl	CCVC MONO→*F,*N -L MONO→*F V _{LONG} (C) MONO→*F SUP→*M,*F	x	x
die Substandardwohnung	-UNG→*M,*N DWELL→*M	Wohnung	-UNG→*M,*N DWELL→*M	✓	✓
die Synärese	-E→*M,*N			✓	✓
die Szientistin	-IN SUFF.→*M,*N WOMEN→*M,*N			✓	✓
die Tagereise	-E→*M,*N	Reise	-E→*M,*N	✓	✓
die Taschenausgabe	-E→*M,*N	Ausgabe	-E→*M,*N	✓	✓
die Teilbarkeit	-HEIT/-(IG)KEIT→*M,*N			✓	✓
die Thematik	-IK→*M,*N			✓	✓
die Toilette	-ETTE→*M,*N			✓	✓
die Trassenführung	-UNG→*M,*N	Führung	-UNG→*M,*N	✓	✓
die Tretmine	-E→*M,*N	Mine	-E→*M,*N	✓	✓
die Trope	-E→*M,*N			✓	✓
die Tschechin	-IN SUFF.→*M,*N WOMEN→*M,*N			✓	✓
die Turbine	-INE→*M,*N			✓	✓
die Tyrierin	-IN SUFF.→*M,*N WOMEN→*M,*N			✓	✓
die Übereile	-E→*M,*N	Eile	-E→*M,*N	✓	✓
die Überlegenheit	-HEIT/-(IG)KEIT→*M,*N			✓	✓
die Überstunde	-E→*M,*N	Stunde	-E→*M,*N UNITS OF TIME→*N	✓	✓

<i>die Umschau</i>	DEVERBAL STEM→*F,*N	<i>Schau</i>	CV MONO→*N DEVERBAL STEM→*F,*N	X	X
<i>die Umwicklung</i>	-UNG→*M,*N	<i>Umwicklung</i>	-UNG→*M,*N	✓	✓
<i>die Unterschreitung</i>	-UNG→*M,*N			✓	✓
<i>die Unversehrtheit</i>	-HEIT/-(IG)KEIT→*M,*N			✓	✓
<i>die Vergessenheit</i>	-HEIT/-(IG)KEIT→*M,*N			✓	✓
<i>die Verleitung</i>	-UNG→*M,*N	<i>Leitung</i>	-UNG→*M,*N	✓	✓
<i>die Verpflegung</i>	-UNG→*M,*N			✓	✓
<i>die Verschlossenheit</i>	-HEIT/-(IG)KEIT→*M,*N			✓	✓
<i>die Verspottung</i>	-UNG→*M,*N			✓	✓
<i>die Vertretungsvollmacht</i>		<i>Macht</i>	-FT/-CHT→*M CVCC MONO→*F	X	X
<i>die Verzögerung</i>	-UNG→*M,*N			✓	✓
<i>die Vindikation</i>	-ION→*M,*N			✓	✓
<i>die Vorherbestimmung</i>	-UNG→*M,*N	<i>Bestimmung</i>	-UNG→*M,*N	✓	✓
<i>die Waagrechte</i>	-E→*M,*N			✓	✓
<i>die Wahlverwandtschaft</i>	-SCHAFT→*M,*N	<i>Verwandtschaft</i>	-SCHAFT→*M,*N	✓	✓
<i>die Weltmarke</i>	-E→*M,*N	<i>Marke</i>	-E→*M,*N	✓	✓
<i>die Wienerin</i>	-IN SUFF.→*M,*N WOMEN→*M,*N			✓	✓
<i>die Wortkargheit</i>	-HEIT/-(IG)KEIT→*M,*N			✓	✓
<i>die Wutrede</i>	-E→*M,*N	<i>Rede</i>	-E→*M,*N	✓	✓
<i>die Zackenkrone</i>	-E→*M,*N	<i>Krone</i>	-E→*M,*N	✓	✓
<i>die Zergliederung</i>	-UNG→*M,*N	<i>Gliederung</i>	-UNG→*M,*N	✓	✓
<i>die Ziellinie</i>	-IE [iə]→*M,*N	<i>Linie</i>	-IE [iə]→*M,*N	✓	✓
<i>die Zoohandlung</i>	-UNG→*M,*N	<i>Handlung</i>	-UNG→*M,*N	✓	✓
<i>die Zufälligkeit</i>	-HEIT/-(IG)KEIT→*M,*N			✓	✓
<i>die Zwangsschiene</i>	-E→*M,*N	<i>Schiene</i>	-E→*M,*N STICK→*N	✓	✓
<i>(das) Antwerpen</i>	CITIES/TOWNS→*M,*F -EN→*F			✓	✓
<i>(das) Bautzen</i>	CITIES/TOWNS→*M,*F -EN→*F			✓	✓
<i>(das) Bessarabien</i>	REGIONS→*M,*F -EN→*F			✓	✓
<i>(das) Großbritannien</i>	COUNTRIES→*M,*F -EN→*F			✓	✓
<i>(das) Karlovy Vary</i>	CITIES/TOWNS→*M,*F			✓	✓
<i>(das) Karlsbad</i>	CITIES/TOWNS→*M,*F	<i>(Bad)</i>	CVC MONO→*F V _{LONG} (C) MONO→*F	✓	✓
<i>(das) Kühlungsborn</i>	CITIES/TOWNS→*M,*F	<i>(Born)</i>	CVCC MONO→*F	✓	✓
<i>(das) Münster</i>	CITIES/TOWNS→*M,*F -ER→*F,*N			X	✓
<i>(das) Nagasaki</i>	CITIES/TOWNS→*M,*F			✓	✓
<i>(das) New Hampshire</i>	REGIONS→*M,*F			✓	✓
<i>(das) Patagonien</i>	REGIONS→*M,*F -EN→*F			✓	✓
<i>(der) Iskariot</i>	NAME: MAN→*F,*N			✓	✓
<i>(der) Klimt</i>	NAME: MAN→*F,*N CCVCC→*F,*N -N(C)(C) MONO→*F,*N			✓	✓
<i>(der) Lärtes</i>	NAME: MAN→*F,*N			✓	✓
<i>(der) Naftali</i>	NAME: MAN→*F,*N			✓	✓
<i>(der) Schiele</i>	NAME: MAN→*F,*N -E→*M,*N			✓	✓

<i>(der) Tintoretto</i>	NAME: MAN → *F, *N -O → *F			✓	✓
<i>(der) Walter</i>	NAME: MAN → *F, *N -ER → *F, *N			✓	✓
<i>(die) Nadja</i>	NAME: WOMAN → *M, *N -A → *M, *N			✓	✓
<i>(die) Rahel</i>	NAME: WOMAN → *M, *N			✓	✓
<i>(die) Sigrun</i>	NAME: WOMAN → *M, *N			✓	✓
<i>(die) Sophia, Sophie</i>	NAME: WOMAN → *M, *N -A → *M, *N			✓	✓
<i>das Außerstreitverfahren</i>	-EN → *F CONVERSION → *M, *F	<i>Verfahren</i>	-EN → *F CONVERSION → *M, *F	✓	✓
<i>das Aviso</i>	-O → *F SEM. EQUIV. → *F, *N (DER HINWEIS)/ → *F (DER/DAS AVIS)			X	X
<i>das Flett</i>	-ETT → *M, *F CCVC MONO → *F, *N			X	✓
<i>das Kulm</i>	-[I]C MONO → *F -N(c)(c) MONO → *F, *N CVCC MONO → *F			X	X
<i>das Sechsfache</i>	-E → *M, *N ADJ. INANIMATE → *M, *F			X	✓
<i>das/der Ar</i>	VC MONO → *M, *F V _{LONG} (C) MONO → *F ∅V _{LONG} MONO → *F SCI. UNITS → *M, *F			✓*	✓*
<i>das/die Korpuskel</i>	SCI. UNITS → *M, *F			✓*	✓*
<i>der Außerstreitrichter</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N	<i>Richter</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N	✓	✓
<i>der Aviso</i>	-O → *F			✓	✓
<i>der Detailhandel</i>		<i>Handel</i>	DEVERBAL STEM → *F, *N	✓	✓
<i>der Farre</i>	MAMMALS → *F DOM. ANIMALS → *F YOUNG ANIMALS → *M, *F MALE ANIMALS → *F, *N -E → *M, *N			✓	✓
<i>der Holländer₁</i>	-ER SUFF. → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Holländer₂</i>	-ER SUFF. → *F, *N CHEESES → *F, *N			✓	✓
<i>der Holländer₃</i>	-ER SUFF. → *F, *N			✓	✓
<i>der Kolm</i> <i>= der/das Kulm</i>	-[I]C MONO → *F -N(c)(c) MONO → *F, *N CVCC MONO → *F			✓*	✓*
<i>der Rahm</i>	-N(c)(c) MONO → *F, *N R- MONO → *F CVC MONO → *F V _{LONG} (C) MONO → *F			✓	✓
<i>der Sigrist</i>	-IST → *F, *N GEN. PERSON → *F, *N			✓	✓
<i>der Tip</i>	T- MONO → *F CVC MONO → *F			✓	✓
<i>der/das Avis</i>	SEM. EQUIV. → *F, *N (DER HINWEIS)/SEM. EQUIV. → *M, *F (DAS AVIS)			✓	✓
<i>der/das Garnknäuel</i>		<i>Knäuel</i>		✓*	✓*
<i>der/das Juchten</i>	-EN → *F (LEATHER TYPES → *M, *F)			✓	✓

<i>der/das New Look</i>	SEM. EQUIV. → *F, *N (DER STIL) / → *M, *F (DAS AUSSEHEN)	<i>Look</i>	CVC MONO → *F SEM. EQUIV. → *F, *N (DER STIL) / → *M, *F (DAS AUSSEHEN)	✓	✓
<i>der/das Renforcé</i>	-EE/-ÉE/-É → *F FABRIC → *F, *N (CONVERSION → *M, *F)			✓*	✓
<i>der/die Kranke</i>	ADJ. MAN → *F, *N/ ADJ. WOMAN → *M, *N -E → *M, *N			✓	✓
<i>der/die Renitente</i>	ADJ. MAN → *F, *N/ ADJ. WOMAN → *M, *N -E → *M, *N			✓	✓
<i>die Angelobung</i>	-UNG → *M, *N			✓	✓
<i>die Elektrische</i>	-E → *M, *N	<i>(Bahn)</i>		✓	✓
<i>die Metropolis</i>	-IS → *M, *N			✓	✓
<i>die Tramway</i>	SEM. EQUIV. → *M, *N (DIE STRAßENBAHN)			✓	✓
<i>Frutti</i>				-	-
<i>Geisa = das Geison</i>	-ON UNSTRESS. → *M, *F			✓	✓
<i>Gerichtskosten</i>				-	-
<i>Korpuskularstrahlen = der Strahl</i>	[f]C- MONO → *F, *N -L MONO → *F V _{LONG} (C) MONO → *F CCCVC MONO → *F, *N STICK → *N (DEVERBAL STEM → *F, *N)			✓	✓

Appendix G: Experiment

Hallo, ich bin Doktorandin am Department of German and Dutch der Universität Cambridge und forsche dort zum Thema „Genus im Deutschen“. Im Rahmen meiner Doktorarbeit führe ich ein kleines Online-Experiment durch, in dem Sie 26 erfundenen Wörtern ein Genus (der, die oder das) zuweisen müssen. Die einzige Voraussetzung dafür ist, dass Sie deutsche/r Muttersprachler/in sind.

Das Experiment wird etwa 10 Minuten Ihrer Zeit in Anspruch nehmen und ich würde mich sehr freuen, wenn Sie mich bei meinen Forschungsarbeiten unterstützen könnten.

Bei Fragen können Sie mich gerne per E-Mail kontaktieren: ec574@cam.ac.uk

Vielen Dank für Ihre Hilfe!

Über folgenden Link gelangen Sie direkt zum Online-Experiment:

https://cambridge.eu.qualtrics.com/jfe/form/SV_6RwG8HDnIGtOyLH

Bitte teilen!

Im Folgenden sehen Sie 26 erfundene Wörter, und für manche Wörter auch eine fiktionale Bedeutung und einen Beispielsatz. Sie müssen jedem Wort ein Genus (*der*, *die* oder *das*) zuweisen. Wenn Sie möchten, können Sie einen Kommentar dazu schreiben – müssen Sie aber nicht. Wenn Sie sich zwischen zwei Genera nicht entscheiden können, dürfen Sie zwei Kästchen ankreuzen. **Es gibt keine richtigen oder falschen Antworten.** Bitte geben Sie einfach Ihre instinktive Reaktion; denken Sie nicht zu lange darüber nach!

Woher kommen Sie? _____ (Land/Bundesland/Region)

Beispiel:

Weulchen

= kleines Meerestier

D___ Weulchen ist nach dem Zwerg-Seepferdchen das zweitkleinste bekannte Meerestier.

- der Weulchen*
- die Weulchen*
- das Weulchen*

Kommentar (optional): _____

1. Zunie

= Stadt in Burkina Faso

z.B. D___ heutige Zunie ist mit 1,6 Millionen Einwohnern die zweitgrößte Stadt des Landes.

- der Zunie*
- die Zunie*
- das Zunie*

Kommentar (optional): _____

2. Yall

= weiblicher Vorname

z.B. Da kommt d___ kluge Yall von gegenüber.

- der Yall*
- die Yall*
- das Yall*

Kommentar (optional): _____

3. Yällchen

= weiblicher Vorname

z.B. D___ kleine Yällchen ist ein gutes Kind.

- der Yällchen*
- die Yällchen*
- das Yällchen*

Kommentar (optional): _____

4. Zabe

= eine Stoffart

z.B. D___ Zabe, woraus das Hemd besteht, ist sehr leicht und gleichzeitig fest.

- der Zabe*
- die Zabe*
- das Zabe*

Kommentar (optional): _____

5. Placht

(es gibt weder eine Definition noch einen Beispielsatz für dieses Wort)

- der Placht*
- die Placht*
- das Placht*

Kommentar (optional): _____

6. Henn

= Maßeinheit in der Elektrodynamik

z.B. Die internationale Einheit der Induktivität ist d___ Henn, benannt nach dem deutschen Physiker Hans Henn.

- der Henn
- die Henn
- das Henn

Kommentar (optional): _____

7. Bazka

= Währung

z.B. D___ Bazka ist die Währung von Mauretanien. 1 Bazka unterteilt sich in 100 Kobo.

- der Bazka
- die Bazka
- das Bazka

Kommentar (optional): _____

8. Fralang

= eingeschossiger Haustyp, der in den Bergregionen in Nepal verbreitet ist

z.B. D___ Fralang der Familie Adhikari wurde in den 60er Jahren errichtet.

- der Fralang
- die Fralang
- das Fralang

Kommentar (optional): _____

9. Klepisma

(es gibt weder eine Definition noch einen Beispielsatz für dieses Wort)

- der Klepisma
- die Klepisma
- das Klepisma

Kommentar (optional): _____

10. Jaff

= Obstsorte aus Südostasien

z.B. D___ Jaff kommt von den indonesischen Inseln Sumatra und Java und wächst in Dickichten auf nassen, sumpfigen Böden.

- der Jaff
- die Jaff
- das Jaff

Kommentar (optional): _____

11. Dranalt

(es gibt weder eine Definition noch einen Beispielsatz für dieses Wort)

- der Dranalt
- die Dranalt
- das Dranalt

Kommentar (optional): _____

12. Aranne

= Fluss in Frankreich

z.B. D___ Aranne verläuft über eine Länge von 203 km und fließt in nordwestlicher Richtung durch das Bergland.

- der Aranne
- die Aranne
- das Aranne

Kommentar (optional): _____

13. Gamie

= achter Buchstabe des georgischen Alphabets (Ⴡ)

z.B. D___ Gamie ist der achte Buchstabe des georgischen Alphabets und kann als Ⴡ oder ႁ geschrieben werden.

- der
- die
- das

Kommentar (optional): _____

14. Springling

= Blume

z.B. D___ Springling wächst aus einer braunen Zwiebel und wird 15 - 40 cm hoch. Die Blüten sind blau oder weiß und haben 6 Blütenblätter.

- der Springling
- die Springling
- das Springling

Kommentar (optional): _____

15. Randor

= Baum

z.B. D___ Randor wächst 20 - 50 m hoch und trägt weit ausladende knorrige Äste.

- der Randor
- die Randor
- das Randor

Kommentar (optional): _____

16. Trelair

(es gibt weder eine Definition noch einen Beispielsatz für dieses Wort)

- der Trelair
- die Trelair
- das Trelair

Kommentar (optional): _____

17. Mellade

= eine Fischart

z.B. D___ Mellade ist eine Fischart aus der Familie der Karpfenfische und kann maximal 80 cm lang und bis zu 7kg schwer werden.

- der Mellade
- die Mellade
- das Mellade

Kommentar (optional): _____

18. Grileau

(es gibt weder eine Definition noch einen Beispielsatz für dieses Wort)

- der Grileau*
- die Grileau*
- das Grileau*

Kommentar (optional): _____

19. Arlt

= Käsesorte aus Zypern

z.B. *D___ Arlt ist ein Schnittkäse und wird aus Schafsmilch hergestellt.*

- der Arlt*
- die Arlt*
- das Arlt*

Kommentar (optional): _____

20. Schnucht

(es gibt weder eine Definition noch einen Beispielsatz für dieses Wort)

- der Schnucht*
- die Schnucht*
- das Schnucht*

Kommentar (optional): _____

21. Vanade

= Name einer Sprache, die in Papua-Neuguinea gesprochen wird

z.B. *D___ Vanade hat die Wortfolge OSV (Objekt-Subjekt -Verb), die sehr selten in den Sprachen der Welt vorkommt.*

- der Vanade*
- die Vanade*
- das Vanade*

Kommentar (optional): _____

22. Ahb

= alkoholisches Getränk aus Thailand

z.B. Je nach Art der Lagerung entwickelt d__ Ahb später einen ganz eigenen Geschmack.

- der Ahb*
- die Ahb*
- das Ahb*

Kommentar (optional): _____

23. Lasp

= eine ursprünglich in den südostasiatischen Tropen und Subtropen beheimatete Insektenart

z.B. D__ Lasp kann gefährliche Viren übertragen, wie das Denguefieber oder das Zika-Virus

- der Lasp*
- die Lasp*
- das Lasp*

Kommentar (optional): _____

24. Syllenage

= Metall

z.B. D__ Syllenage kann aus Erdöl durch Bildung einer Emulsion unter Zugabe von Wasser und Magnesiumnitrate extrahiert werden.

- der Syllenage*
- die Syllenage*
- das Syllenage*

Kommentar (optional): _____

25. Felpt

= Spiel

z.B. D__ Felpt wird heute im Mittleren Osten auch Trick-Track genannt.

- der Felpt*
- die Felpt*
- das Felpt*

Kommentar (optional): _____

26. Grett

(es gibt weder eine Definition noch einen Beispielsatz für dieses Wort)

- der Grett*
- die Grett*
- das Grett*

Kommentar (optional): _____

Vielen Dank für Ihre Hilfe!