



## **Discours**

Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics

**26 | 2020**  
**Varia**

---

# Pronoun Processing and Interpretation by L2 Learners of Italian: Perspectives from Cognitive Modelling

Petra Hendriks and Margreet Vogelzang

---



### **Electronic version**

URL: <http://journals.openedition.org/discours/10679>

DOI: 10.4000/discours.10679

ISSN: 1963-1723

### **Publisher:**

Laboratoire LATTICE, Presses universitaires de Caen

### **Electronic reference**

Petra Hendriks and Margreet Vogelzang, « Pronoun Processing and Interpretation by L2 Learners of Italian: Perspectives from Cognitive Modelling », *Discours* [Online], 26 | 2020, Online since 02 November 2020, connection on 03 November 2020. URL : <http://journals.openedition.org/discours/10679> ; DOI : <https://doi.org/10.4000/discours.10679>

---





Revue de linguistique, psycholinguistique et informatique

<http://journals.openedition.org/discours/>

## **Pronoun Processing and Interpretation by L2 Learners of Italian: Perspectives from Cognitive Modelling**

---

Petra Hendriks

Center for Language and Cognition Groningen  
University of Groningen

Margreet Vogelzang

Department of Theoretical and Applied Linguistics  
University of Cambridge

.....  
Petra Hendriks, Margreet Vogelzang, « Pronoun Processing and Interpretation by L2 Learners of Italian: Perspectives from Cognitive Modelling », *Discours* [En ligne], 26 | 2020, mis en ligne le 2 novembre 2020.

.....  
URL : <http://journals.openedition.org/discours/10679>

.....  
Titre du numéro : *Varia*

Coordination : Nicolas Hernandez & Lydia-Mai Ho-Dac

Date de réception de l'article : 16/12/2019

Date d'acceptation de l'article : 25/03/2020



Presses  
universitaires  
de Caen 



# Pronoun Processing and Interpretation by L2 Learners of Italian: Perspectives from Cognitive Modelling

---

Petra Hendriks

Center for Language and Cognition Groningen  
University of Groningen

Margreet Vogelzang

Department of Theoretical and Applied Linguistics  
University of Cambridge

.....

How do second language learners acquire form-meaning associations in the second language that are inconsistent with their first language? In this study, we focus on subject pronouns in Italian and Dutch. A native speaker of the non-null subject language Dutch learning the null subject language Italian as a second language will not only have to learn to use and comprehend null pronouns, but will also have to learn to use and comprehend overt pronouns differently in the L2 than in the L1. The interpretation of Italian overt pronouns, but not of Dutch overt pronouns or Italian null pronouns, has been argued to require perspective taking, specifically the use of hypotheses about the conversational partner's communicative choices to guide one's own choices. Therefore, a related question is how perspective taking and cognitive constraints influence L2 acquisition of such forms. Using computational cognitive modelling, this study explores two learning scenarios. In cognitive model 1, second language acquisition proceeds in the same way as first language acquisition and is based on the same grammar. In cognitive model 2, second language acquisition differs from first language acquisition and involves the construction of a partly different grammar. Our results suggest that the second scenario may be cognitively more plausible than the first one. Furthermore, our models explain why second language learners of Italian perform less native-like on overt pronouns than on null pronouns.

**Keywords:** cognitive modelling, null pronouns, overt pronouns, Italian, second language acquisition

## 1. Introduction

- 1 One of the major puzzles in linguistic research is why second language learners often fail to achieve full native competence in their second language, as is evidenced by differences in their linguistic performance compared to that of native speakers. An influential hypothesis on second language acquisition is the Interface Hypothesis (Sorace & Filiaci, 2006; Sorace, 2011). According to this hypothesis, acquiring properties of the second language that reside in one linguistic domain (such as syntax or phonology) is not problematic, but acquiring properties of the second language that lie at the interface between syntax and another cognitive domain (in particular the interface between syntax and pragmatics) is difficult. This hypothesis was inspired

by the finding that, in a test of subject pronoun interpretation, highly proficient, near-native speakers of L2 Italian gave different responses from monolingual native Italian speakers with respect to overt pronouns (Sorace & Filiaci, 2006). This contrasted with their interpretation of null subjects, which was native-like.

2 A relevant question is why the interface between syntax and pragmatics is particularly difficult for L2 speakers. Is there a persistent influence of the learners' L1 on the L2 (as Tsimplici et al. [2004] suggest) and does the linguistic knowledge of L2 speakers therefore differ from that of L1 speakers? Or do L2 learners use different processing strategies to process interface structures? Arguing for the second type of explanation, Sorace (2011) states that L2 speakers are less efficient in their processing and integration of information from different domains in real time than monolingual speakers. However, differences in linguistic knowledge and differences in processing strategies are likely not independent. So how do linguistic knowledge and sentence processing interact in acquiring a second language?

3 Pronouns have often been used as a case study for investigating the acquisition of a second language. They may be particularly challenging to master at a native-like level, as pronominal reference is acquired "late" even by monolinguals and requires both language-internal (e.g., syntax) and language-external (e.g., working memory) resources and components to be developed and applied (Tsimplici, 2014). Indeed, it is known that bilinguals and second language learners struggle to become native-like in their interpretation (Serratrice, 2007; Sorace & Filiaci, 2006) as well as in their production (e.g., Serratrice et al., 2004; Torregrossa et al., 2019) of pronouns.

4 In this study we focus on subject pronouns in Italian and Dutch. A native speaker of Dutch learning Italian as a second language will not only have to learn to use and comprehend null pronouns, which occur in Italian but do not occur in Dutch, but will also have to learn to use and comprehend overt pronouns differently in the L2 than in the L1. How do differences in the pronominal systems of Dutch and Italian affect the L2 acquisition of Italian by L1 Dutch speakers? And how do cognitive processes influence the L2 acquisition of pronouns? Using computational cognitive modelling, which allows us to investigate the interplay between linguistic and cognitive constraints on sentence processing, this study aims to shed more light on these issues.

## 2. Theoretical background

5 In this section, we provide the theoretical background of our study. Section 2.1 discusses the pronominal systems of Dutch and Italian and highlights the relevant differences. In Section 2.2, we introduce the method of computational cognitive modelling and the specific cognitive architecture used in this study, ACT-R. Section 2.3 discusses two earlier cognitive modelling studies on monolingual pronoun processing, one on Dutch and one on Italian, on which we base our models of L2 acquisition.

## 2.1. Pronoun use and interpretation

- 6 In languages such as English and Dutch, overt pronouns are used to refer to entities that are prominent in the linguistic discourse (e.g., Ariel, 1990; Givón, 1983). For example, in the English sentence pair “Yesterday John went to the movies. He had a great time”, the overt subject pronoun *he* in the second sentence refers to the individual John mentioned in the first sentence. The referent John is prominent in the linguistic discourse, because he was referred to in the preceding sentence. Another factor influencing prominence is syntactic role (Arnold, 1998; Torregrossa et al., 2015). In the sentence pair “Yesterday John went to the movies with his brother. He had a great time”, most native speakers of English will again interpret *he* as referring to John. This is because *John* in the first sentence is the grammatical subject, whereas *his brother* is a prepositional object. Because of its syntactic prominence, *John* is the most likely antecedent of the pronoun. If a referent is less prominent, for example John’s brother in the previous example, most native speakers will not use a pronoun to refer to this referent, but instead be more explicit and use a full noun phrase (Arnold, 1998; Givón, 1983). That is, they will say “Yesterday John went to the movies with his brother. His brother had a great time”.
- 7 Languages such as Italian, Spanish, Catalan, Romanian, and Arabic have a pronominal system that differs from that of English and Dutch. These languages are called null subject languages because, in addition to overt pronouns, they also allow null pronouns as subjects. Whereas in English and Dutch there is a division of labour between overt pronouns and full noun phrases, with overt pronouns generally referring to the most prominent referent in the linguistic discourse and full noun phrases referring to other referents, in null subject languages we see a similar division of labour between null pronouns and overt pronouns. Null pronouns are generally used to refer to the most prominent referent in the linguistic discourse, and overt pronouns are generally used to refer to a less prominent referent (cf. Carminati, 2002; Filiaci, 2010; Tsimpli et al., 2004; Vogelzang et al., 2020).
- 8 For English and Dutch, it has been argued that the choice between an overt pronoun and a full noun phrase is not merely based on the preceding linguistic discourse. Rather, referential choice has been argued to also require the speaker to consider the perspective of the listener and to estimate whether the listener will be able to retrieve the intended referent on the basis of the form used (Hendriks et al., 2008 and 2014). Suppose the speaker would like to use a pronoun. Placing herself in the perspective of the listener, however, the speaker could arrive at the conclusion that the listener will not be able to retrieve the intended referent for this pronoun, for example because the intended referent is not the most prominent referent in the discourse. As a consequence, the speaker should then block the pronoun and use a more informative full noun phrase instead. This illustrates the process of perspective taking, by which we mean the use of hypotheses about the conversational partner’s communicative choices to guide one’s own communicative choices.

9 Evidence for the involvement of perspective taking in mature referential choice comes from children's overuse of pronouns to refer to less-prominent referents (Hendriks et al., 2014). Further evidence comes from the observation of a correlation between children's use of full noun phrases for less-prominent referents and their theory of mind abilities, that is, their understanding that other people have beliefs and intentions that are different from one's own (Kuijper et al., 2015).

10 Likely, speakers of Italian face a similar task when deciding between a null pronoun and an overt pronoun. That is, a speaker of Italian intending to use a null pronoun might conclude – based on the linguistic discourse and the process of perspective taking – that the listener will not be able to retrieve the intended referent of the null pronoun in the given discourse, and may thus decide to use an overt pronoun instead.

11 This process of perspective taking may not only be relevant in pronoun production, but also in pronoun comprehension. Vogelzang (Vogelzang, 2017; Vogelzang et al., 2020) argues that null pronouns in Italian are grammatically encoded to refer to the most prominent referent (i.e., the discourse topic), but that overt pronouns in Italian may be ambiguous in their reference. To disambiguate overt pronouns, the listener will have to reason about the referential choice of the speaker and whether the speaker would have used an overt pronoun to express the intended meaning. This process of perspective taking is expected to be cognitively demanding. Indeed, using pupil size as a measure of cognitive effort, Vogelzang et al. (2020) found that in monolingual Italian adults the increase in pupil size was smaller for null pronouns than for overt pronouns. This indicates that, in Italian, null pronouns are less costly to process than overt pronouns. A relevant question is how this process of perspective taking interacts with other factors in the acquisition of reference in a second language. To investigate this issue, we use the method of cognitive modelling.

## 2.2. Cognitive modelling

12 A relatively novel way to study the effects of cognitive constraints on human behaviour is through cognitive modelling. Cognitive modelling involves the development of computational models within a cognitive architecture. A cognitive architecture is an integrated theory of cognition that aims to describe and explain universal aspects of the human cognitive system, and at the same time is a piece of software that can be executed to yield actual performance. By developing models within a cognitive architecture, the models are guaranteed to satisfy the particular set of theoretical premises and assumptions the cognitive architecture is built on, which are derived from multiple experiments on human behaviour. Cognitive models additionally include the required knowledge to perform a particular task. They computationally simulate human behaviour on this particular task and generate predictions about human performance that can be empirically tested. Thus, the performance of cognitive models on a task can be compared to actual human performance on the same task. Similarities and differences between the output of the cognitive models and the output of human participants will provide insights into the way cognitive abilities

such as memory influence human performance on a task, for example resolving a pronoun in the native language or a second language.

- 13 In this study, we make use of the cognitive architecture ACT-R (Adaptive Control of Thought – Rational; Anderson, 2007; Anderson et al., 2004). ACT-R aims to explain human cognition as the interaction between functional modules. The ACT-R modules that are the most relevant for language processing are declarative memory and procedural memory. Declarative memory stores factual information as chunks. This factual information can be used if the chunk is retrieved from declarative memory. The higher the activation value of the chunk, the more likely it is to be retrieved and the faster it will be retrieved. Procedural memory consists of production rules (if-then rules) that are executed one by one. If several production rules meet the requirements for being executed, the production rule with the highest utility will be selected. Important properties of ACT-R are that all processing is serial and that each operation takes time. If production rules are often used in succession, they can be merged into a new production rule, making processing more efficient and faster (for details see Taatgen & Lee, 2003). Thus, experience results in higher processing speed.
- 14 In the context of language learning, experience equals language exposure: more language exposure results in higher processing speed. Further factors involved in language learning are the linguistic properties of the language to be learned, the linguistic knowledge the learner already possesses, the cognitive abilities of the language learner, and the cognitive constraints on language learning. The cognitive abilities of the language learner and the cognitive constraints on learning follow from the properties of ACT-R. In our cognitive models, these cognitive abilities and constraints interact with linguistic knowledge and linguistic processes in a way that we will discuss below. Other factors that have been argued to be relevant for second language acquisition, such as social factors (e.g., language status, situational features) and affective factors (e.g., attitudes, motivation, anxiety), will not be taken into account in this study.
- 15 Our cognitive models have in common with usage-based theories that first and second language acquisition are assumed to be influenced by general cognitive mechanisms and amount of exposure to particular forms (see Ellis & Cadierno, 2009, for an overview of the factors involved in usage-based accounts of second language acquisition). However, our cognitive models differ from usage-based theories in other aspects, for example in the use of formal-linguistic constraints and the relation between forms and meanings. Usage-based theories assume that language is made up of conventionalised form-meaning pairings, called constructions (see also Goldberg, 2003), in which form and meaning are inextricably related (Ellis & Cadierno, 2009). In contrast, in our cognitive models the mapping from meaning to form in production and the mapping from form to meaning in comprehension are two distinct processes. As a consequence, language learners may relate a form to a particular meaning in production but not in comprehension, for example with pronouns, as we will see below. Through the interplay between formal-linguistic and cognitive constraints, form-meaning pairings gradually become established during language acquisition.

### 2.3. Earlier cognitive models of pronoun processing

16 In this study, we investigate how language exposure interacts with linguistic and cognitive constraints when learning a second language. By second language we mean a language that is acquired after having acquired (substantial parts of) the first language. The second language in this study is Italian, which is a null subject language allowing for null subject pronouns in addition to overt subject pronouns. The first language in this study is Dutch, which is a non-null subject language very similar to English as far as subject pronouns are concerned. Focusing on the language pair Dutch-Italian allows us to benefit from the insights already obtained from the cognitive models developed to account for the use and interpretation of overt subject pronouns by monolingual Dutch adults and children (Van Rij et al., 2013) and the use and interpretation of null and overt subject pronouns by monolingual Italian adults and children (Vogelzang, 2017).

17 The cognitive models of Van Rij et al. (2013) and Vogelzang (2017) consist of two components. The first component is the discourse component. The discourse component keeps track of the referents in the linguistic discourse. If a referent is mentioned in the linguistic discourse, the model either adds a new referent to its discourse representation, or updates the information in memory about an existing referent. Each referent is represented as a chunk in declarative memory and has a certain activation value. The more often and the more recently a referent is mentioned, the higher its activation value. On top of this, the subject of the clause receives an extra activation boost to model the effects of syntactic prominence. The activation values of the various referents thus change over time. The referent with the highest activation value at a given moment is considered to be the most prominent referent and is identified as the discourse topic. However, due to the noise that is inherently present in ACT-R models, there is a possibility that sometimes a less prominent referent is accidentally identified as the discourse topic.

18 The second component of the cognitive models of Van Rij et al. (2013) and Vogelzang (2017) is the sentence processing component. Sentence processing takes place after the discourse component has identified the discourse topic. In the sentence processing component, linguistic constraints are implemented as chunks in declarative memory. These linguistic constraints are taken from linguistic accounts of pronoun use and interpretation (e.g., Hendriks et al., 2008; Vogelzang, 2017) that are cast in the constraint-based linguistic framework of Optimality Theory (OT – Prince & Smolensky, 2004). The sentence processing component can be used for both sentence production and sentence comprehension. To use a chunk for production or comprehension, for example a referring expression, a discourse referent, or a linguistic constraint, the chunk must first be retrieved from declarative memory by an ACT-R production rule. Since all operations in ACT-R take time, the application of a production rule also takes time. However, the more often the production rule is applied, the faster this process will be.

19 In sentence production as well as in sentence comprehension, the cognitive models incorporate perspective taking in the sentence processing component by

combining the two modalities in a two-step process. To comprehend a sentence, the models start with a comprehension step and select a particular meaning for the encountered form. This is followed by a production step to check whether the selected meaning is indeed the meaning giving rise to the encountered form in production. Vice versa, to produce a sentence the models start with a production step, followed by a comprehension step to check whether the intended meaning is indeed recoverable from the selected form. Since two steps take more time than one step, perspective taking is expected to take additional time. The more often the relevant forms and meanings have been encountered, the more often the relevant production rules were applied and the more likely it is that perspective taking is successfully completed.

20 In our study, we focus on sentence comprehension and consider two different cognitive models for learning pronoun meanings in a second language. In the first cognitive model, we assume that second language learning proceeds in the same way as first language acquisition and is based on the same set of linguistic constraints. In the second cognitive model, we assume that second language learning differs from first language acquisition and may involve the construction of a partly different grammar consisting of a partly different set of linguistic constraints.

### 3. Cognitive model 1

21 In our first cognitive model, model 1, we simulate the situation in which language learners start out with a universal set of linguistic constraints. In the linguistic framework of Optimality Theory, first language acquisition is hypothesised to proceed by adjusting the ranking of the constraints on the basis of positive evidence (e.g., Boersma & Hayes, 2001; Tesar & Smolensky, 1998). Theoretical analysis of the use and interpretation of pronouns in the two languages (see Appendix) suggests that the different patterns in Dutch and Italian are not characterised by a different ranking of the same constraints. Rather, differences seem to arise as a result of the different available anaphoric subject forms in the two languages. Therefore, in model 1 we adopt as a starting point for second language acquisition the grammar and constraint ranking of the first language. For the interpretation of subject pronouns, this means that learners with Dutch as their L1 learning Italian as an L2 apply their linguistic constraints for Dutch to Italian null pronouns and thus gain experience in applying the corresponding production rules. This is expected to result in changes in the activation of the production rules that may lead to native-like performance in Italian. The details of this approach will be explained below.

#### 3.1. Characteristics and assumptions of the model

22 Model 1 contains the linguistic constraints on the use and interpretation of subject pronouns as implemented in ACT-R by Van Rij et al. (2013) for Dutch (the constraints are numbered by us for ease of exposition):

REFECON1: Avoid full noun phrases

REFECON2: Avoid overt pronouns

PROTOP1: Overt pronouns refer to the discourse topic

23 The first two constraints, REFECON1 and REFECON2, are part of a family of constraints promoting referential economy and punishing the use of linguistic material to refer to discourse entities. REFECON1 being stronger than REFECON2, it is more important to avoid full noun phrases than it is to avoid overt pronouns. Hence, overt pronouns are preferred over full noun phrases. The third constraint, PROTOP1, restricts the interpretation of overt pronouns to reference to the discourse topic. It is violated by any occurrence of an overt pronoun that does not refer to the discourse topic. Hendriks et al. (2008) argue that, in the grammar of Dutch, the constraints are ranked in this order, with REFECON1 and REFECON2 dominating PROTOP1.

24 As follows from the way these constraints are phrased, REFECON1 and REFECON2 (and in fact the entire family of referential economy constraints) apply in production only, since the choice between two forms is only relevant in production. In contrast, PROTOP1 applies in comprehension, when the optimal meaning must be determined for an overt pronoun.

25 In addition to the three constraints mentioned above, model 1 also contains the following constraint that is argued by Vogelzang (2017) to be relevant for Italian:

PROTOP2: Null pronouns refer to the discourse topic

26 In model 1, which is the model implemented to account for subject pronouns in Dutch, the constraint PROTOP2 initially is the weakest of the four constraints, as Dutch does not have null pronouns. Note that, although Dutch does not have null pronouns like Italian, Dutch subjects may be omitted if they are topics. This is called topic drop and is not restricted to subjects (see, e.g., Broekhuis & Corver, 2016, where it is pointed out that in speech topic drop is more frequent with objects than with subjects).

27 These constraints interact to determine the optimal form for an input meaning (in production) and the optimal meaning for an input form (in comprehension) (see Appendix). In production, if a speaker of Dutch wishes to refer to the discourse topic, the speaker can either use an overt pronoun or a full noun phrase. Using an overt pronoun satisfies the constraint REFECON1 (which punishes the use of full noun phrases). On the other hand, this referential choice violates the constraint REFECON2 (which punishes the use of overt pronouns). Since using a full noun phrase violates the strongest constraint REFECON1, whereas using an overt pronoun only violates the weaker constraint REFECON2, an overt pronoun satisfies the total set of constraints better than a full noun phrase. Hence, an overt pronoun is the optimal form for expressing reference to the discourse topic.

28 If a speaker wishes to refer to a non-topic, an overt pronoun again is the optimal form. Again, a full noun phrase violates the strongest constraint REFECON1 (punishing the use of full noun phrases). So despite the fact that using an overt

pronoun violates the weaker constraint REF<sub>ECON2</sub>, an overt pronoun is the optimal form to refer to a non-topic. Indeed, this appears to be the pattern shown by Dutch-speaking children, who overuse overt pronouns and also produce pronouns to refer to referents that are not prominent in the discourse (see Hendriks et al., 2014). As is argued by Hendriks et al. (2008), mature speakers of Dutch take an additional step in production and also consider the opposite communicative perspective of the listener. This allows them to check whether the intended meaning is recoverable by listeners from the form selected by the speaker. Taking the perspective of the listener, a mature speaker of Dutch will be able to conclude that an overt pronoun will be interpreted as the topic, which is different from the input meaning (a non-topic). Hence, this form must be blocked for expressing a non-topic meaning and a more explicit form must be used, which is a full noun phrase. As a consequence, mature speakers of Dutch use a full noun phrase to refer to a non-topic.

29 In comprehension, the listener's task is to determine the meaning of the form heard. Here, we focus on whether a pronoun refers to the discourse topic or to another referent. Recall that the constraints REF<sub>ECON1</sub> and REF<sub>ECON2</sub> are only relevant in production and do not play a role in comprehension. Therefore, PRO<sub>TOP1</sub> determines how Dutch pronouns are comprehended. If a listener speaking Dutch encounters a Dutch sentence with an overt subject pronoun, selecting reference to the discourse topic as its meaning satisfies the strongest constraint PRO<sub>TOP1</sub>, whereas selecting reference to a non-topic meaning violates this constraint. As none of the other constraints are relevant for selecting the optimal meaning, the optimal meaning for an overt pronoun is reference to the discourse topic, in accordance with PRO<sub>TOP1</sub>.

30 The constraints of the grammar are implemented in the cognitive model as chunks, which must be retrieved from memory by production rules. The more often these constraints are used, the faster they are retrieved from memory. Thus, language experience through exposure to the relevant linguistic forms and meanings results in the model becoming faster and more efficient in processing these forms and meanings. When the model receives training in Dutch and is presented with overt pronouns and discourse referents, the model retrieves the chunks and uses the production rules for the first three constraints (REF<sub>ECON1</sub>, REF<sub>ECON2</sub>, PRO<sub>TOP1</sub>). Since Dutch does not have null pronouns, the constraint PRO<sub>TOP2</sub> is not used by the model. Hence, the chunk for constraint PRO<sub>TOP2</sub> and its corresponding production rules are still weak and are only retrieved from memory slowly.

31 Italian, as opposed to Dutch, has null pronouns. As a consequence, a speaker wishing to refer to the topic will use a null pronoun, as this form is the referentially most economical form in the language and hence satisfies the constraints best. For reference to a non-topic, a null pronoun also is the optimal form according to the constraints of the grammar. Therefore, this situation is similar to the situation in Dutch with overt pronouns. As Vogelzang (2017) argues, this is the pattern shown by Italian-speaking children, since their processing speed is still low. In contrast, the processing speed of mature speakers of Italian is sufficiently high to take into

account the perspective of the listener. Hence, they block null pronouns for reference to a non-topic because the listener will interpret a null pronoun as the topic, and will produce an overt pronoun instead.

32 In comprehension, a native Italian listener encountering a null pronoun will interpret this null pronoun as referring to the topic, in accordance with PRO<sub>TOP2</sub>. In contrast, an overt pronoun will be found ambiguous between a topic and a non-topic interpretation. To assign a mature non-topic interpretation to overt pronouns, listeners will have to consider the perspective of the speaker and block the topic meaning, as this meaning is expressed by speakers by using a null pronoun.

33 In model 1, the constraints and their ranking are the same for the L1 and the L2. In addition, the activation mechanism for discourse referents in the model is also the same in the L1 and the L2, and no differentiation is made between the syntactic structures of the two languages. Although this is a simplification, we believe this does not affect our modelling results since we focus on subject pronouns in one sentence type.

34 Now what would a mature Dutch listener do when encountering a null pronoun, as in Italian? Let us assume that the Dutch listener initially uses his or her Dutch grammar and Dutch constraint ranking to interpret this null pronoun. According to the constraint PRO<sub>TOP2</sub>, null pronouns refer to the discourse topic. At the same time, however, in accordance with the constraint PRO<sub>TOP1</sub>, according to the Dutch grammar overt pronouns also refer to the discourse topic. So a Dutch listener is predicted to interpret all encountered pronouns as referring to the discourse topic, irrespective of whether these are overt or null pronouns. The question we address using cognitive modelling is whether, and if so how, Dutch listeners learn to interpret Italian null and overt pronouns in a native-like way on the basis of linguistic experience only, given the cognitive constraints that are part of the cognitive model and given the linguistic assumptions specified above.

### 3.2. Language input to the model

35 Acquiring a language is only possible if the learner receives input in that language. As input in the second language, the cognitive model receives Italian sentences similar to the ones used in Tsimplì et al. (2004) and Sorace and Filiaci (2006). Each sentence consists of two clauses connected by a subordinate conjunction. The first clause contains a definite noun phrase as the subject and a definite noun phrase as the direct object, thus introducing two discourse referents. The second clause contains an overt subject pronoun (*lui* [he]) or a null subject pronoun ( $\emptyset$ ) that could be interpreted as referring to one of these two discourse referents. An example of an Italian sentence used in our study is [1].

[1] L'uomo saluta il ragazzo, mentre lui/ $\emptyset$  attraversa la strada  
 the man greets the boy while he/ $\emptyset$  crosses the street  
 'the man greets the boy, while he/ $\emptyset$  crosses the street'

36 The corresponding sentences in Dutch, but with only overt subject pronouns (*bij* [he]) as the subject of the second clause, were used for training in and testing on the first language, Dutch. See [2] for an example of a Dutch sentence.

[2] De man groet de jongen, terwijl hij de straat oversteekt  
 the man greets the boy while he the street crosses  
 ‘the man greets the boy, while he crosses the street’

37 Ten different sentences in total were used for testing on each language. To obtain a sufficient number of sentences for testing purposes, per test phase the sentences were repeated several times. To prevent a learning effect taking place during testing as a result of the repetition of the test sentences, the discourse model was erased before each repetition.

### 3.3. Training and testing of the model

38 To provide a starting point for second language acquisition, the cognitive model is first trained on the first language, Dutch. The training data for the first language consist of overt pronouns. The task of the model is to determine for each pronoun whether it refers to the topic or to a non-topic. This procedure provides the model with experience in processing the linguistic constraints in the L1. In total, the model receives 2,000 training items. This phase simulates the acquisition and end state of the first language. Note that the term “training” in this study does not refer to explicit and conscious training, but rather to the language experience children receive and base their knowledge of their native language on. In this training phase, as in later training phases, the model does not receive explicit feedback on the target interpretation of the pronouns.

39 In the second phase, the model is tested on the L1 Dutch. This allows us to establish the model’s performance on overt pronoun comprehension in the L1 and compare its performance to the performance of native speakers of Dutch. Different runs of the model simulate different participants; in total in our study the model simulates 20 different participants. Due to the noise in the cognitive model, performance differs slightly per simulated participant. The simulated participants all receive 40 test sentences containing an overt pronoun. The model’s task in the test is to determine the referent of this pronoun, which could be either the referent referred to by the subject of the preceding main clause (that is, the topic), or the referent referred to by the object of the main clause (that is, a non-topic).

40 Next, the model is tested on overt and null pronoun comprehension in Italian, and is then trained and tested on Italian pronoun comprehension twice. This is done to simulate a language learner (1) at the start of the process of second language acquisition without having received any input in the second language yet, (2) after having received some input in the second language, and (3) after having received a substantial amount of input in the second language. Again, the term “training” must be understood in an implicit and unconscious sense. The Italian items in the

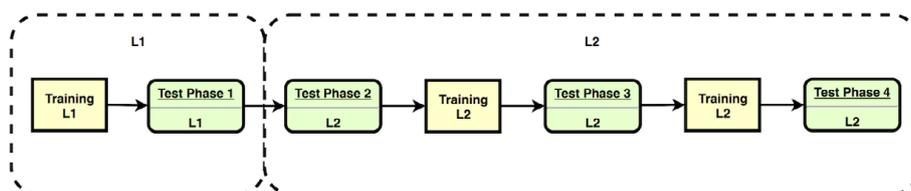


Figure 1 – Overview of the training phases (yellow) and test phases (green) in the L1 (Dutch, left) and the L2 (Italian, right)

training input contain pronouns in a proportion of 30% overt pronouns and 70% null pronouns (cf. Serratrice et al., 2004, and Lorusso et al., 2005, who report that 68% and 74%, respectively, of subjects were null pronouns in a sample of child-directed speech; unfortunately, as far as we know no percentages are available for the speech specifically directed to second language learners). The model's task in the test is to determine the referent of the null or overt pronoun.

41 The different phases of the cognitive modelling procedure are listed below and are graphically represented in Figure 1.

Phase 1: Training in L1 (2,000 trials)

Phase 2: Testing on L1 (40 sentences, 20 simulated participants) – test 1 (Dutch)

Phase 3: Testing on L2 (40 sentences, 20 simulated participants) – test 2 (Italian)

Phase 4: Training in L2 (1,000 trials)

Phase 5: Testing on L2 (40 sentences, 20 simulated participants) – test 3 (Italian)

Phase 6: Training in L2 (1,000 trials)

Phase 7: Testing in L2 (40 sentences, 20 simulated participants) – test 4 (Italian)

42 In the next section, the results of the tests performed with this cognitive model are discussed.

#### 4. Results and discussion of model 1

43 Figure 2 (test 1) shows the model's performance on pronoun comprehension in the L1 Dutch. In our test materials, when the subordinate clause is encountered, the discourse topic is the entity referred to by the subject of the main clause. This is because this referent is the syntactically most prominent referent referred to by the previous clause. Therefore, selection of the subject of the main clause as the referent of the pronoun in the subordinate clause indicates reference to the discourse topic, whereas selection of the object of the main clause as the referent of the pronoun in the subordinate clause indicates reference to a non-topic. The model generally interprets the overt subject pronoun in the subordinate clause as referring to the subject of the main clause. That is, the pronoun *hij* [he] in the subordinate clause in [2] is taken to refer to the man in the large majority of cases (89%), and to the boy in only a minority of cases (the remaining 11%). This corresponds to the results of Vogelzang et al. (2016), who found that adult native speakers of Dutch select the subject of the main clause as the referent of an overt subject pronoun in 91% of cases.

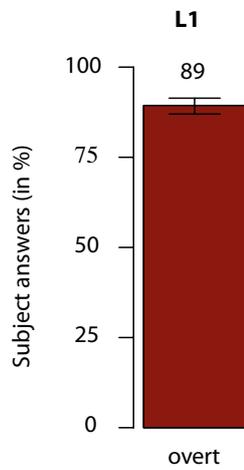


Figure 2 – Pronoun comprehension in L1 (Dutch, test 1)

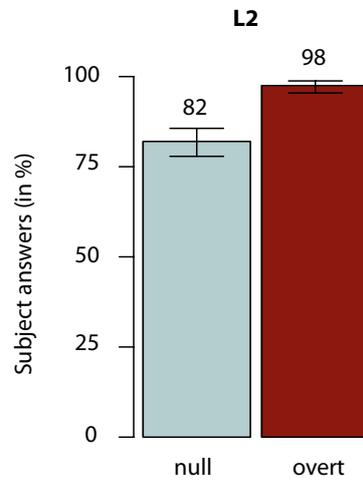


Figure 3 – Pronoun comprehension in L2 without training (Italian, test 2)

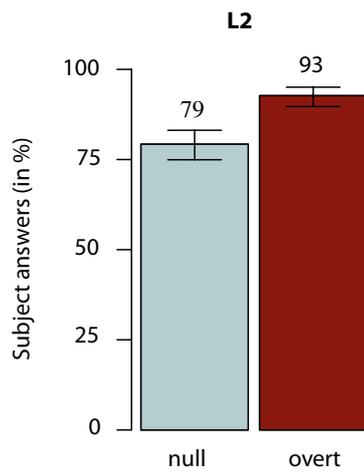


Figure 4 – Pronoun comprehension in L2 after one training phase (Italian, test 3)

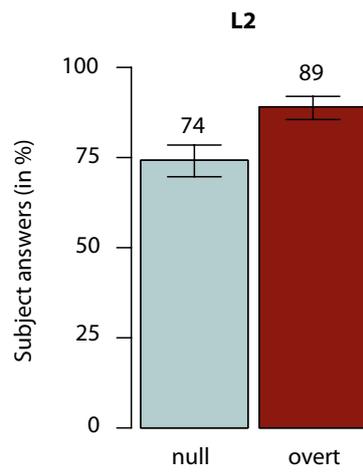


Figure 5 – Pronoun comprehension in L2 after two training phases (Italian, test 4)

44 Figures 3, 4 and 5 show the model’s performance on pronoun comprehension in the L2 Italian after different amounts of training.

45 Without any training in Italian (Figure 3, test 2), the model interprets Italian overt pronouns as referring to the subject of the main clause in almost all cases (98%). This is different from the way overt pronouns are interpreted by native speakers of Italian; in a study by Vogelzang et al. (2020) with sentences like [1] in a story context, adult native speakers of Italian provided 39% subject responses for overt pronouns. Null pronouns, which were not encountered before by the model because Dutch does not have null

pronouns, are nevertheless interpreted by the model as referring to the subject in most cases (82%). This corresponds to the way the native speakers of Italian in the study by Vogelzang et al. (2020) interpreted null pronouns, providing 86% subject responses. The results of Tsimpli et al. (2004) with native Italian controls on sentences like [1], but in a task different from that of Vogelzang et al. (2020), show the same pattern of interpretations for overt and null pronouns, contrasting with the results of model 1. However, their participants gave lower percentages of subject responses overall: 7.6% subject responses for overt pronouns, and 50.75% subject responses for null pronouns.

46 After one training phase in Italian (Figure 4, test 3), the percentage of interpretations of overt pronouns as referring to the subject decreases (93%). In addition, the percentage of interpretations of null pronouns as referring to the subject also decreases slightly (79%). A second training phase in Italian (Figure 5, test 4) strengthens this pattern. This developmental pattern is even clearer in Figure 6, which presents the same data as in Figures 3, 4, and 5, but now ordered by amount of training.

47 The model results at the different test times were analysed using generalised linear mixed-effects regression models (“lme4” package in R, Bates et al., 2014). Models were constructed for null and overt pronouns separately, each time with interpretation (subject or non-subject) as the dependent variable, test time as a fixed factor, and intercepts for simulated participants as a random factor. For null pronouns, a significant decrease in subject answers was observed over time (i.e., over tests;  $\beta = -0.01$ ;  $z = -2.66$ ;  $p < 0.01$ ). For overt pronouns, a significant decrease in subject answers was observed as well ( $\beta = -0.01$ ;  $z = -4.51$ ;  $p < 0.001$ ). This decrease reflects that over time the model more often selects reference to a non-topic (non-subject). However, the model has not yet adequately learned in which situations non-topic reference is preferred, as even after a substantial amount of input in the second language (test 4) overt pronouns are still interpreted as referring to the subject 89% of the time.

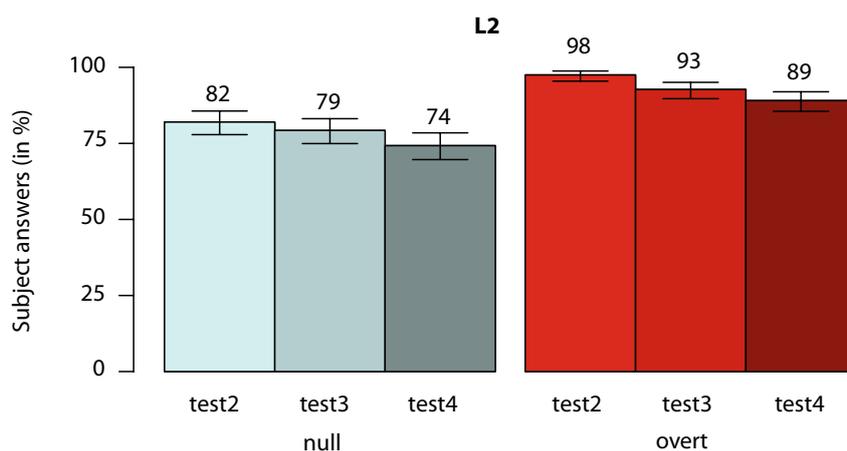


Figure 6 – Comprehension of null pronouns (left) and overt pronouns (right) in L2 (Italian) without training (test 2), after one training phase (1,000 items, test 3) and after two training phases (2,000 items, test 4) in model 1

48 These results indicate that reference of overt pronouns to the topic is difficult to un-learn. In Dutch, overt pronouns refer to the topic, but a speaker of Dutch learning Italian as a second language needs to un-learn this association for Italian, as in Italian overt pronouns do not refer to the topic. The observation that this association is difficult to un-learn makes sense in the model, since model 1 includes the constraint *PROTOP1* stating that overt pronouns refer to the topic, which is relevant in Dutch. The production rules for this constraint must be deactivated in order to obtain the Italian pattern of interpretation. Without any indication that an overt pronoun should not refer to the topic, however, this constraint may continue to be used for the resolution of overt pronouns by Dutch learners of Italian.

49 In contrast, reference of null pronouns to the topic is easy to learn, for two reasons. First of all, the constraint is already present in the grammar as part of the universal constraint hierarchy (see Section 3.1). And secondly, reference to the topic is the meaning most frequently selected by the model in the past. Therefore, it will be selected most often by the model. This follows from the cognitive constraints in the model: the more frequent the meaning, the easier it will be for the model to retrieve it from declarative memory and process it by applying the relevant production rules. In addition, because the production rules for reference to the topic have been used most often in the past (namely in the L1), they will be faster and have a larger probability of being selected.

50 One could argue that perhaps more training is necessary to achieve native-like performance on pronoun comprehension in a second language. To investigate this, we also modelled the process of second language acquisition with 10,000 rather than 1,000 items per training phase in the L2. Despite the much larger amount of language input in Italian (with 10 times as many training items in the second language Italian as in the first language Dutch), the developmental pattern remains largely the same. That is, the model still has the tendency to select the topic as the referent of an overt pronoun.

51 In summary, this first series of model simulations shows that without explicit feedback, reference of overt pronouns to the topic is difficult to un-learn in an L2 when the L1 contains similar overt pronouns that refer to the topic. Recall that the Dutch grammar and Dutch constraint ranking were used for interpreting Italian as well; all constraints were shared between the two languages. However, it has also been argued that second language learners develop separate grammars (either almost completely separated, or interconnected to a lesser or greater degree, see Cook [2002] for discussion). In the following sections, where we consider a second series of model simulations, we investigate this option.

## 5. Cognitive model 2

52 In our second model, model 2, we simulate the situation in which language learners start out with a partially distinct set of linguistic constraints for the second language.

In this situation, the learner's task when learning a second language is to learn which constraints, that may or may not be identical to those in the first language, are relevant for the grammar of the second language. For subject pronouns, the task of learners with Dutch as their L1 learning Italian as an L2 is therefore to learn that Italian does not have a constraint on the use and interpretation of overt pronouns.

### 5.1. Characteristics and assumptions of the model

53 Model 2 includes the following linguistic constraints on the use and interpretation of subject pronouns, which were also part of model 1:

REFECON1: Avoid full noun phrases

REFECON2: Avoid overt pronouns

PROTOP2: Null pronouns refer to the discourse topic

54 The only difference between model 2 and model 1 is that, in contrast to model 1, in model 2 the grammar does not contain the constraint PROTOP1, which specifies that overt pronouns refer to the discourse topic.

55 Removal of a constraint from the grammar is not compatible with the learning mechanisms proposed in Optimality Theory for first language acquisition. Rather, it suggests different learning mechanisms for second language acquisition and for first language acquisition. In model 2, removal of PROTOP1 was implemented as an immediate decision of the model. One way to account for this decision is through explicit learning: the second language learner may have received explicit instructions about the use and interpretation of overt pronouns. This possibility is in line with the view advocated in usage-based linguistics (e.g., Wulff & Ellis, 2018) that, whereas first language acquisition depends primarily on implicit learning, second language acquisition is largely characterised by explicit learning. Another way to account for this decision is through implicit learning: based on the language input in the second language, the second language learner may realise that PROTOP1 pertaining to overt pronouns does not hold for the second language. How exactly the learning process is motivated, whether it is an explicit or implicit process, and whether it is possible to implement this learning process as a process of gradual development rather than as an immediate decision is left for further study. We will return to this issue in the general discussion.

### 5.2. Language input to the model

56 The language input to model 2 is the same as the language input to model 1, see Section 3.2.

### 5.3. Training and testing of the model

57 Model 2 is trained and tested in the same way as model 1, see Section 3.3.

## 6. Results and discussion of model 2

58 Figure 7 presents the performance of model 2 ordered by amount of training.

59 The model results at the different test times were again analysed using generalised linear mixed-effects regression models (see Section 4 for more details). For null pronouns, a decrease in subject answers was observed over time, but this effect was not significant ( $\beta = -0.01$ ;  $z = -1.38$ ;  $p = 0.17$ ). For overt pronouns, no effects were seen either ( $\beta = 0.01$ ;  $z = 0.19$ ;  $p = 0.85$ ). These results thus indicate that the beneficial effects of not using the constraint PROTOP1 for the interpretation of overt pronouns is instant, without much learning taking place in additional training sessions.

60 So, without the constraint PROTOP1 on overt pronouns, the cognitive model performs much more native-like even without much training. Performance of the model on null pronouns as well as on overt pronouns is similar to the results of Vogelzang et al. (2020), who found 86% subject responses for null pronouns and 39% subject responses for overt pronouns in a study with adult native speakers of Italian. However, although the general pattern of interpretations is similar, the model shows more subject interpretations for overt pronouns than the native speakers in the study of Vogelzang et al. did. This is in line with the results of Sorace and Filiaci (2006; English second language learners of Italian) and Serratrice (2007; 6- to 9-year-old English-Italian bilingual children), who both found more subject interpretations for overt pronouns for second language learners or bilingual children compared to their monolingual peers, despite showing native-like interpretations of null pronouns. Model 2 thus seems to capture pronoun processing as well as the pattern of non-native interpretations displayed by second language learners in experiments.

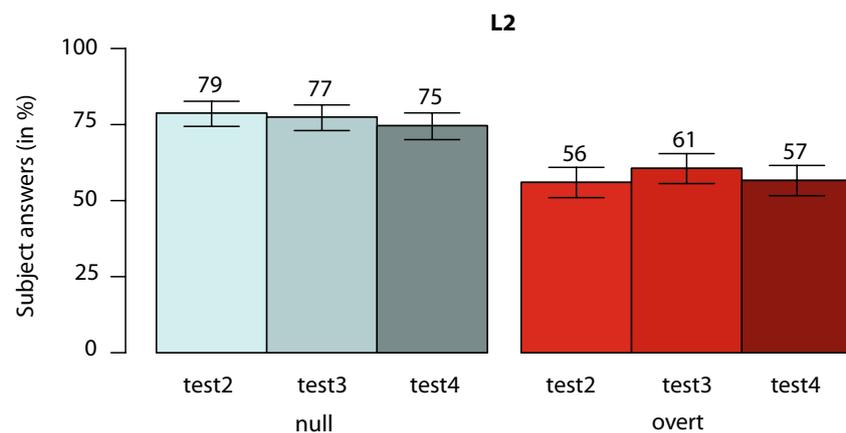


Figure 7 – Comprehension of null pronouns (left) and overt pronouns (right) in L2 (Italian) without training (test 2), after one training phase (1,000 items, test 3) and after two training phases (2,000 items, test 4) in model 2

## 7. General discussion

61 In this study, we addressed the question of how differences in the use and interpretation of pronouns between Dutch and Italian affect the acquisition of pronouns in Italian as a second language by native speakers of Dutch. In addition, we wanted to know what role perspective taking and cognitive constraints play in the acquisition of pronouns in the second language. Using computational cognitive modelling, two learning scenarios were explored.

62 In cognitive model 1, second language acquisition was modelled in the same way as first language acquisition and was based on the same grammar. This grammar included the constraint *ProTop1*, which requires overt pronouns to refer to the topic. This constraint reflects the way overt pronouns are interpreted in the L1 Dutch, but is inconsistent with the way overt pronouns are interpreted in the L2 Italian. Our results show that, based on our linguistic and cognitive assumptions, the association in Dutch between overt pronouns and reference to the topic is difficult to un-learn in the L2 Italian. Even after a substantial amount of input in the L2, overt pronouns in the L2 were still interpreted as referring to the topic most of the time, due to the presence of the constraint *ProTop1*. This suggests that un-learning the association between a particular form and meaning in the L1 by weakening the relevant processing rules (i.e., production rules in ACT-R) and constraints (i.e., chunk activation in ACT-R) is not a cognitively plausible scenario for attaining native-like competence in a second language.

63 In cognitive model 2, second language acquisition was modelled differently from first language acquisition and involved the construction of a partly different grammar. Like model 1, this grammar included the constraints on referential economy and the constraint *ProTop2* requiring null pronouns to refer to the topic, but did not include the constraint *ProTop1* from the L1 requiring overt pronouns to refer to the topic. Without the constraint *ProTop1*, the model performed much more native-like, even without much training. This finding suggests that second language learners need to construct a grammar for the second language that is at least partially different from that of the first language to be able to learn aspects of the grammar that are different in the two languages. To learn these aspects, it is not sufficient that the processing rules for L1 constraints that are not consistent with associations in the L2 merely become weaker in the second language. Based on cognitively plausible assumptions (as incorporated in the cognitive architecture ACT-R) and linguistic assumptions (which are supported by empirical evidence from various studies), our models instead suggest that these inconsistent constraints need to be deactivated, so that the grammar and processing rules of the second language must be partly different from the grammar and processing rules of the first language.

64 This finding seems to be in line with conclusions of usage-based accounts, according to which learning a second language amounts to learning a new set of form-meaning mappings (Ellis & Cadierno, 2009). Consistent with usage-based accounts, this study also finds that second language acquisition is different from,

and more complex than, first language acquisition because knowledge of the L2 is in competition with knowledge of the L1. However, note that the assumptions on which our conclusions are based are partly different from the assumptions of usage-based linguistics (see Section 2.2).

65 The learning strategy following from cognitive model 2 starts out with a grammar for the second language that could be maximally identical to the grammar of the first language, but deactivates constraints if there is evidence that these constraints are not relevant for the second language. This conclusion presents us with a learning puzzle: How can the second language learner learn that the constraint PROTOP1 is not relevant for the second language on the basis of positive evidence only? In model 2, we implemented removal of this constraint as an immediate decision, but clearly for the account to be cognitively plausible this constraint needs to be removed on the basis of evidence available to the learner. Earlier, we suggested two options: through explicit instructions about the use and interpretation of overt pronouns, or through implicit learning. Implicit learning requires that the language learner encounters examples where overt pronouns in Italian do not refer to the discourse topic. These examples are not infrequent, so this may provide us with a plausible account of how PROTOP1 is un-learned. An example of the relevant learning evidence could be formed by the following Italian sentence taken from Sorace and Filiaci (2006), in which the overt subject pronoun *lei* [she] refers to the object and non-topic Giovanna, and not to the subject and topic Mario: *Quando Mario chiama Giovanna, lei è contenta* [when Mario calls Giovanna, she is happy]. However, only if we simulate the learning process on the basis of this type of evidence through cognitive modelling is it possible to determine whether this potential account is compatible with our cognitive assumptions.

66 In our study, we modelled the interpretation of overt pronouns in Italian, in contrast to the interpretation of null pronouns in Italian and overt pronouns in Dutch, as a process requiring the listener to consider the perspective of the speaker. This accounts for the pattern of interpretation of Italian pronouns. In particular, it explains the division of labour between null and overt pronouns in Italian, according to which null pronouns generally refer to topics, whereas overt pronouns refer to less-prominent referents. In our cognitive models, perspective taking is modelled as a two-step process requiring sufficient processing speed and language experience to be completed successfully. If perspective taking is not completed successfully, overt pronouns are interpreted as referring to the discourse topic. Since, in our cognitive models, processing overt pronouns is more complex than processing null pronouns, which do not rely on perspective taking for their interpretation, our models predict that second language learners perform less native-like on overt pronouns than on null pronouns. Indeed, the second language learners and bilingual learners in the experiments by Sorace and Filiaci (2006) and Serratrice (2007) tended to assign more topic interpretations to overt pronouns than monolinguals do. Thus, our cognitive models are able to predict and capture the pattern of non-native interpretations displayed by second language learners in experiments.

67 To conclude, we investigated the acquisition of overt and null pronoun interpretation in L2 Italian, using the method of computational cognitive modelling. This method allowed us to investigate the effects of linguistic and cognitive constraints on pronoun interpretation in a second language, giving equal weight to the formal-linguistic account of pronoun interpretation and to its interactions with general cognition. This shows the potential of cognitive modelling to investigate further issues in second language acquisition, such as individual differences in L2 interpretation and processing based on, for example, variation in working memory capacity, or the various stages in second language development. Ultimately, this method may provide a “cognitively grounded” theory of linguistic competence in a second language.

## References

- ANDERSON, J.R. 2007. *How Can the Human Mind Occur in the Physical Universe?* Oxford – New York: Oxford University Press.
- ANDERSON, J.R., BOTHELL, D., BYRNE, M.D., DOUGLASS, S., LEBIERE, C. & QIN, Y. 2004. An Integrated Theory of the Mind. *Psychological Review* 111 (4): 1036-1060.
- ARIEL, M. 1990. *Accessing Noun-Phrase Antecedents*. Londres – New York: Routledge.
- ARNOLD, J.E. 1998. *Reference Form and Discourse Patterns*. PhD thesis. Stanford University.
- BATES, D.M., MAECHLER, M., BOLKER, B. & WALKER, S. 2014. *Package “lme4”: Linear Mixed-Effects Models using “Eigen” and S4*. R package version 1.1-19. Online: <http://cran.r-project.org/web/packages/lme4/lme4.pdf> (retrieved on 6 November 2018).
- BOERSMA, P. & HAYES, B. 2001. Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32 (1): 45-86.
- BROEKHUIS, H. & CORVER, N. 2016. *Syntax of Dutch: Verbs and Verb Phrases*. Amsterdam: Amsterdam University Press. Vol. 3.
- CARMINATI, M.N. 2002. *The Processing of Italian Subject Pronouns*. PhD thesis. University of Massachusetts at Amherst.
- COOK, V.J. (ed.) 2002. *Portraits of the L2 User*. Clevedon – Buffalo – Toronto: Multilingual Matters.
- ELLIS, N.C. & CADIerno, T. 2009. Constructing a Second Language. Introduction to the Special Section. *Annual Review of Cognitive Linguistics* 7: 111-139.
- FILIACI, F. 2010. *Anaphoric Preferences of Null and Overt Subjects in Italian and Spanish: A Cross-Linguistic Comparison*. PhD thesis. University of Edinburgh.
- GIVÓN, T. (ed.) 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam – Philadelphia: J. Benjamins.
- GOLDBERG, A.E. 2003. Constructions: A New Theoretical Approach to Language. *Trends in Cognitive Science* 7 (5): 219-224.
- HENDRIKS, P., ENGLERT, C., WUBS, E. & HOEKS, J.C.J. 2008. Age Differences in Adults’ Use of Referring Expressions. *Journal of Logic, Language and Information* 17 (4): 443-466.

- HENDRIKS, P., KOSTER, C. & HOEKS, J.C.J. 2014. Referential Choice across the Lifespan: Why Children and Elderly Adults Produce Ambiguous Pronouns. *Language, Cognition and Neuroscience* 29 (4): 391-407.
- KUIJPER, S.J.M., HARTMAN, C.A. & HENDRIKS, P. 2015. Who Is He? Children with ASD and ADHD Take the Listener into Account in their Production of Ambiguous Pronouns. *PLoS ONE* 10 (7): 1-18. Available online: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0132408>.
- LORUSSO, P., CAPRIN, C. & GUASTI, M.T. 2005. Overt Subject Distribution in Early Italian Children. In A. BRUGOS, M.R. CLARK-COTTON & S. HA (eds.), *BUCLD 29 Online Proceedings Supplement*. Boston: Boston University: 1-7. Available online: <https://www.bu.edu/buclid/files/2011/05/29-LorussoBUCLD2004.pdf>.
- PRINCE, A. & SMOLENSKY, P. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden – Oxford – Victoria: Blackwell Publishing.
- SERRATRICE, L. 2007. Cross-Linguistic Influence in the Interpretation of Anaphoric and Cataphoric Pronouns in English-Italian Bilingual Children. *Bilingualism: Language and Cognition* 10 (3): 225-238.
- SERRATRICE, L., SORACE, A. & PAOLI, S. 2004. Crosslinguistic Influence at the Syntax-Pragmatics Interface: Subjects and Objects in English-Italian Bilingual and Monolingual Acquisition. *Bilingualism: Language and Cognition* 7 (3): 183-205.
- SORACE, A. 2011. Pinning Down the Concept of “Interface” in Bilingualism. *Linguistic Approaches to Bilingualism* 1 (1): 1-33.
- SORACE, A. & FILIACI, F. 2006. Anaphora Resolution in Near-Native Speakers of Italian. *Second Language Research* 22 (3): 339-368.
- TAATGEN, N.A. & LEE, F.J. 2003. Production Compilation: A Simple Mechanism to Model Complex Skill Acquisition. *Human Factors* 45 (1): 61-76.
- TESAR, B. & SMOLENSKY, P. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29 (2): 229-268.
- TORREGROSSA, J., BONGARTZ, C. & TSIMPLI, I.M. 2015. Testing Accessibility: A Cross-Linguistic Comparison of the Syntax of Referring Expressions. *Linguistic Society of America: Annual Meeting Extended Abstracts* 6: 1-4. Available online: <http://journals.linguisticsociety.org/proceedings/index.php/ExtendedAbs/article/download/3046/2791>.
- TORREGROSSA, J., BONGARTZ, C. & TSIMPLI, I.M. 2019. Bilingual Reference Production. A Cognitive-Computational Account. *Linguistic Approaches to Bilingualism* 9 (4-5): 569-599.
- TSIMPLI, I.M. 2014. Early, Late or Very Late? Timing Acquisition and Bilingualism. *Linguistic Approaches to Bilingualism* 4 (3): 283-313.
- TSIMPLI, I.M., SORACE, A., HEYCOCK, C. & FILIACI, F. 2004. First Language Attrition and Syntactic Subjects: A Study of Greek and Italian Near-Native Speakers of English. *International Journal of Bilingualism* 8 (3): 257-277.
- VAN RIJ, J., VAN RIJN, H. & HENDRIKS, P. 2013. How WM Load Influences Linguistic Processing in Adults: A Computational Model of Pronoun Interpretation in Discourse. *Topics in Cognitive Science* 5 (3): 564-580.

- VOGELZANG, M. 2017. *Reference and Cognition: Experimental and Computational Cognitive Modeling Studies on Reference Processing in Dutch and Italian*. PhD thesis. University of Groningen.
- VOGELZANG, M., FOPPOLO, F., GUAISTI, M.T., VAN RIJN, H. & HENDRIKS, P. 2020. Reasoning about Alternative Forms Is Costly: The Processing of Null and Overt Pronouns in Italian Using Pupillary Responses. *Discourse Processes* 57 (2): 158-183.
- VOGELZANG, M., HENDRIKS, P. & VAN RIJN, H. 2016. Pupillary Responses Reflect Ambiguity Resolution in Pronoun Processing. *Language, Cognition and Neuroscience* 31 (7): 876-885.
- WULFF, S. & ELLIS, N.C. 2018. Usage-Based Approaches to Second Language Acquisition. In D. MILLER, F. BAYRAM, J. ROTHMAN & L. SERRATRICE (eds.), *Bilingual Cognition and Language. The State of the Science across Its Subfields*. Amsterdam – Philadelphia: J. Benjamins: 37-56.

## Appendix

This appendix presents the analyses in Optimality Theory supporting the theoretical accounts discussed in Section 3.1. The device used in Optimality Theory to show how constraint interaction yields a particular output is called a *tableau*. In a tableau, constraints are listed in the top row from left to right in order of decreasing strength. The input of optimisation is provided in the top cell of the first column. Possible outputs are listed below the input. Constraint violations for a particular candidate output are marked with an asterisk (\*) in the corresponding cell. A fatal constraint violation (i.e., one that makes the candidate output suboptimal) is marked with an exclamation mark (!). Grey cells indicate that the corresponding constraint does not apply. The pointing finger (☞) indicates the optimal candidate, which is the output of optimisation.

Dutch (cf. Hendriks et al., 2007; Hendriks et al., 2014)

Input: Topic	REFECON1	REFECON2	ProTOP1
☞ Overt pronoun		*	
Full noun phrase	* !		

*Tableau 1* – Production of a topic in Dutch.

Considering the perspective of the listener (*Tableau 3*) does not change the output

Input: Non-topic	REFECON1	REFECON2	ProTOP1
☞ Overt pronoun		*	
Full noun phrase	* !		

*Tableau 2* – Production of a non-topic in Dutch.

Considering the perspective of the listener (*Tableau 3*) results in blocking an overt pronoun for a non-topic meaning and selecting a full noun phrase instead

Input: Overt pronoun	REFECON1	REFECON2	ProTOP1
☞ Topic			
Non-topic			* !

*Tableau 3* – Comprehension of an overt pronoun in Dutch.

Considering the perspective of the speaker (*Tableau 1*) does not change the output

	Input: Full noun phrase	REFECON1	REFECON2	ProTop1
☞	Topic			
☞	Non-topic			

*Tableau 4* – Comprehension of a full noun phrase in Dutch.

Considering the perspective of the speaker (*Tableaux 1 and 2*) results in blocking a topic meaning for a full noun phrase and selecting a non-topic meaning

Note that adding the constraint ProTop2 (null pronouns refer to the discourse topic) to the grammar of Dutch would not change the pattern shown in *Tableaux 1-4*, as ProTop2 does not apply to candidate outputs without a null pronoun.

Italian (cf. Vogelzang, 2017)

	Input: Topic	REFECON1	REFECON2	ProTop2
☞	Null pronoun			
	Overt pronoun		* !	
	Full noun phrase	* !		

*Tableau 5* – Production of a topic in Italian.

Considering the perspective of the listener (*Tableau 7*) does not change the output

	Input: Non-topic	REFECON1	REFECON2	ProTop2
☞	Null pronoun			
	Overt pronoun		* !	
	Full noun phrase	* !		

*Tableau 6* – Production of a non-topic in Italian.

Considering the perspective of the listener (*Tableau 7*) results in blocking a null pronoun for a non-topic meaning and selecting an overt pronoun instead

	Input: Null pronoun	REFECON1	REFECON2	ProTop2
☞	Topic			
	Non-topic			* !

*Tableau 7* – Comprehension of a null pronoun in Italian.

Considering the perspective of the speaker (*Tableau 5*) does not change the output

	Input: Overt pronoun	REFECON1	REFECON2	ProTOP2
☞	Topic			
☞	Non-topic			

*Tableau 8* – Comprehension of an overt pronoun in Italian.  
 Considering the perspective of the speaker (*Tableaux 5 and 6*) results in blocking a topic meaning for an overt pronoun and selecting a non-topic meaning

	Input: Full noun phrase	REFECON1	REFECON2	ProTOP2
☞	Topic			
☞	Non-topic			

*Tableau 9* – Comprehension of a full noun phrase in Italian.  
 Considering the perspective of the speaker (*Tableaux 5 and 6*) results in blocking a topic meaning for a full noun phrase and selecting a non-topic meaning

Note that adding constraint ProTOP1 (overt pronouns refer to the discourse topic) to the grammar of Italian would change the pattern of interpretation of overt pronouns. Overt pronouns would then incorrectly be interpreted as the topic in all cases. However, as in *Tableau 8*, considering the perspective of the speaker would result in the mature Italian pattern by blocking a topic meaning and selecting a non-topic meaning instead.