Spanish and Greek subjects in contact: Greek as a heritage language in Chile

Aretousa Giannakou

Magdalene College

University of Cambridge

May 2018

This dissertation is submitted for the degree of Doctor of Philosophy

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

It does not exceed the prescribed word limit for the relevant Degree Committee.

Aretousa Giannakou

May 2018

Cambridge, UK

Abstract

The present study aims to capture linguistic variation in subject distribution of two typologically similar languages, Greek and Chilean Spanish, considering adult monolingual and bilingual speakers of Greek as a heritage/minority language in Chile. The focus is on null and overt third-person subjects in topic-continuity and topic-shift contexts. Such structures involve the interface between syntax and discourse/pragmatics, a vulnerable domain in bilingualism. Previous research has shown overextension of the scope of the overt subject pronoun in contexts where null subjects are discursively expected (e.g. Tsimpli, Sorace, Heycock & Filiaci 2004). The Interface Hypothesis (IH) (Sorace 2011) was formulated to account for such findings, which obtain even in pairs of two null subject languages (Sorace, Serratrice, Filiaci & Baldo 2009). The key question as to the language-contact effects on subject distribution in pairs of two null subject languages requires further exploration while the combination of Greek and Spanish has been so far understudied. The IH is evaluated with new empirical data from a bilingual situation not studied before.

Data from oral narratives and aural pronominal anaphora resolution were elicited from monolinguals and three types of bilinguals, namely first-generation immigrants, heritage speakers and L2 speakers of Greek residing in Chile. The monolingual data revealed differences in the use and interpretation of overt subject pronouns between Greek and Chilean Spanish. The crosslinguistic difference lies in the strong deictic properties of the Greek pronoun compared to its Spanish counterpart; hence differences obtain because of the relative strength of the two pronominal forms. No overextension of the scope of overt pronouns was found in bilinguals, against predictions stemming from the Interface Hypothesis. This may relate to the typological similarity between Greek and Spanish as well as to the nature of the Greek pronoun, which makes its use relatively categorical. Such findings lend support to the Representational account (Tsimpli et al. 2004). On the contrary, null subjects gave rise to optionality presumably due to their complexity, which demands higher degrees of computational efficiency. The Vulnerability Hypothesis (Prada Pérez 2018) may also account for the findings.

Acknowledgements

Contents

3. Research framework

4. Greek and Spanish studies

Part II

5. The study

6. Study 1: Narratives

7. Study 2: Anaphora Resolution

8. General Discussion

9. Conclusions

References

Appendices

List of Tables

List of Figures

Acronyms and abbreviations

Agr	agreement
ANOVA	Analysis of Variance
AoB	age of onset of bilingualism
AP	antecedent preference(s)
AR	anaphora resolution
CI	confidence interval
DP	determiner phrase
EPP	Extended Projection Principle
HS	heritage speaker(s)
IH	Interface hypothesis
IMM	first-generation immigrants
L1	first language
L2	second language / second language speaker(s)
L2er(s)	second language speaker(s)
LoR	length of residence in the host country
LS	lexical subject(s)
N	number
NS	null subject(s)
n.s.	non-significant
NSP	null subject parameter
OR	odds ratios
OSP	overt subject pronoun(s)
PAH	Position of antecedent hypothesis
RAE	Real Academia Española
RC	relative clause(s)
RR	relative risk
SD	standard deviation
SE	standard error
Spec	specifier
TAM	tense, aspect, mood
TC	topic continuity
TP	tense phrase
TS	topic shift
UG	Universal Grammar
VH	Vulnerability hypothesis
YoB	number of years of bilingualism

But I, the Mind, [...] move among the phenomena which I create,
I distinguish between them conveniently; I unite them with laws
and yoke them to my heavy practical needs.
I impose order on disorder
and give a face - my face - to chaos.

Nikos Kazantzakis
Ascesis: Saviors of God
(1922/1945)

Part I

1. Introduction

1.1 Overview

The present study is situated within the broad framework of adult bilingualism as a phenomenon of language contact in the context of migration. It considers the case of Greek in contact with Spanish in Chile focusing on the morphosyntactic and discourse-pragmatic properties of subject expression and interpretation.

By Greek, reference is made to the contemporary Greek spoken in Greece, where 'the vast majority of Greek speakers now speak a common language with only relatively minor dialectal variations' (Holton, Mackridge & Philippaki-Warburton 2012: xxi). Spanish is a Romance language with official status in several countries including most countries of Latin America. Two main 'standard' norms are recognised, namely Peninsular and Latin American Spanish, while there are also regional varieties. Chilean Spanish is the variety spoken in Chile and it is classified as a separate dialect zone (Lipski 1994, 2012). It is relatively homogeneous showing very little regional variation only at the vernacular level (Lipski 1994; Moreno Fernández 2014). The distinctive features of this dialect vis-à-vis other varieties of Spanish are mainly lexical and phonological. The present study involves monolingualism in Greek and Chilean Spanish and it also explores Greek as spoken in Chile by Greek-Spanish bilinguals.

Bilingualism refers to speakers possessing knowledge of two languages and the ability to use them even if these are not mastered equally and without regard to proficiency levels (e.g. Baker 2011; Genesee 2016). Bilingual speakers generally have a stronger and a weaker language, the former being more native-like or dominant than the latter (Montrul 2008). They are specific speakers-hearers who develop competence in each language to the extent required by their needs and those of

1

their environment (Grosjean 2008). Bilinguals are typically classified according to order, age and context of language acquisition. Simultaneous bilinguals are exposed to two languages from birth, whereas successive/sequential bilinguals learn the second language (L2) after the acquisition of the first language (L1). In particular, adult late bilinguals learn the L2 after the age of 15 or post-puberty (Montrul 2008; Pavlenko 2014; Sorace 2016).

Once a speaker becomes sequential bilingual, there is 'some degree of traffic from L2 to L1' (Schmid 2011: 12). Bilingual situations may thus involve L1 attrition, arising after an 'extensive and intensive period of language contact' (Montrul 2005: 200; see also Tsimpli, Sorace, Heycock & Filiaci 2004). Co-activation of languages, transfer or disuse result in prior linguistic knowledge being less accessible or even modified because of the presence of the new language, which affects L1 production, processing or comprehension (Schmid & Köpke 2017a). L1 attrition occurs at the individual level and may be circumstantial, temporary or permanent (Seliger 1996; Montrul 2008, 2017; Schmid 2011; Chamorro & Sorace 2018). The severity of attrition depends on diverse interacting factors, the most important being the regularity of L1 use (frequency and recency) as well as the length of L2 exposure and use (Kaltsa, Tsimpli & Rothman 2015; Schmid & Köpke 2017a). L1 attrition effects in adult sequential bilinguals are usually minimal and, in particular, L1 core syntactic aspects, such as macroparametric properties, are generally well preserved (e.g. Tsimpli et al. 2004; Montrul 2008; Iverson 2012; Kaltsa et al. 2015; Yılmaz & Schmid 2018). Less efficient processing is what selectively affects certain areas and particularly those involving the interface between syntax and discourse/pragmatics (e.g. Tsimpli et al. 2004; Sorace 2005, 2016; Chamorro & Sorace 2018).

If bilinguals acquire a minority language as an L1 in childhood but subsequently become dominant in the majority language, they are heritage speakers (HS)[1] (Benmamoun, Montrul & Polinsky 2013). According to Montrul (2016b: 2), 'heritage languages are commonly spoken by immigrants and their children', hence are described as 'culturally or ethnolinguistically minority languages that develop in a

---

[1] Typically, HS are dominant in the majority language but they can also be balanced or dominant in the heritage language, although this is unusual.

bilingual setting where another socio-politically majority language is spoken' (see also Rothman 2009b; Polinsky 2018). Due to the shift to the socially-dominant language, which entails reduced input in the heritage language, HS do not usually reach full development of the latter. Inaccessibility to formal education in the heritage language, which is generally common, contributes to this phenomenon. The insufficient input may be also qualitatively different since it is mostly provided by first-generation immigrants whose L1/heritage grammar may have undergone attrition. Thus, the ultimate attainment of many adult HS is the result of incomplete/differential acquisition (Montrul 2016b; cf. Kupisch & Rothman 2016) or divergent attainment (Polinsky 2018). Since HS acquire the heritage language naturalistically in early childhood, they qualify as native speakers of both their languages (Rothman & Treffers-Daller 2014; Montrul 2016b).

Monolingualism, on the other hand, refers to speakers having a single native language and marginal or no knowledge of other languages. Bilinguals generally differ from monolinguals due to crosslinguistic competition and lower degrees of language activation, implying additional cognitive load for language processing (Sorace 2011, 2016; Schmid & Köpke 2017a; Yılmaz & Schmid 2018). In current accounts, comparing bilinguals to monolinguals is sometimes considered to be simplistic (Grosjean 2008; Genesee 2016; Kupisch & Rothman 2016; cf. Montrul 2016b). This is because bilingual grammars are recognised as different from monolingual grammars. The two languages do not exist in isolation as autonomous linguistic systems, therefore cannot be regarded as a combination of two encapsulated systems (Montrul 2008). Comparing monolingual and bilingual linguistic behaviour may, however, be insightful in cases of language pairs, such as Greek and Spanish, which are underexplored as to particular linguistic phenomena. Moreover, comparisons between monolinguals and different types of bilinguals may illuminate potential factors contributing to the (non-)variable linguistic behaviour observed in bilingualism.

Greek as a minority or heritage language is relatively understudied in bilingual situations especially with a host language other than English (but see Kaltsa et al. 2015) and in contact with Latin American Spanish (but see Zombolou 2011). Greek as

spoken in Chile mostly relates to three groups of speakers: (a) first-generation immigrants, considered to be potential L1 attriters; (b) simultaneous/early bilinguals (HS); and (c) L2 speakers (L2ers) who have a strong personal or family connection with the language. HS are said to 'straddle the boundaries between first and second language acquisition' (Benmamoun, Montrul & Polinsky 2010: 14). In the context of the present study, the L2ers seem to straddle the boundaries between heritage and ordinary L2 language acquisition due to the type of connection of these speakers with the Greek language. More precisely, these speakers have close family bonds with Greeks, i.e. they are grandchildren of Greek immigrants or married to Greeks with children speaking heritage Greek. Some of them had also been immigrants in Greece. The three bilingual groups constitute a diverse population with different language exposure patterns and acquisition trajectories. Nevertheless, they all have Greek-Spanish bilingualism as a common denominator in a monolingual country where Chilean Spanish is the dominant language, sharing the situation of their two languages being in asymmetric societal contact. These groups were a priori distinguished due to significant differences in their bilingual profiles.

Considering Greek and Spanish morphosyntax, subject distribution is an interesting ground to explore since both languages are consistent null subject (NS) languages. Subject realisation and interpretation concern anaphora or reference to discourse entities, which establish crucial coherence relations enabling the creation of meaning (Zulaica-Hernández & Gelormini-Lezama 2017). 'If human cognition is fundamentally intentional in the sense of being about or directed towards something, reference is a form of shared intentionality in which the cognitive focus of two or more persons is aligned and jointly focused' (Sidnell & Enfield 2017: 217). In particular, reference as determined by subjects or topics, i.e. discourse-old elements in the common ground of interlocutors in subject position, is the focus of the present study.

The study investigates production and interpretation of third-person null and overt subjects in Greek in the different subpopulations of Greek-Chilean Spanish bilinguals. Adult monolingual-native performance in Greek and Chilean Spanish is first explored comparatively. The two languages are typologically similar and generally pattern alike in subject distribution. The focus on subject structures concerns two discourse

contexts: topic continuity (TC) and topic shift (TS). As NS languages, Greek and Spanish have more options for expressing subjects than non-NS languages, i.e. NS, lexical subjects (LS) or overt subject pronouns (OSP). The choice between the available referential forms is not arbitrary but regulated by discourse-pragmatics.

Different studies in Greek and Spanish suggest the possibility of crosslinguistic variation in the scope of pronominal subjects in the two languages. The first aim of the study was to explore crosslinguistic differences in this area between Greek and Chilean Spanish. Subject distribution in NS languages is regarded as an unstable area in bilingualism due to syntax-discourse/pragmatics interface conditions (e.g. Tsimpli et al. 2004; Sorace 2011; Kaltsa et al. 2015). In certain cases, such instability is due to crosslinguistic influence, while in other cases it seems to be triggered by more general effects of bilingualism. In any case, it typically surfaces as indeterminacy at the level of discourse/pragmatics in the distribution of pronominal forms with the scope of OSP being particularly affected as predicted by the Interface Hypothesis (IH) (Sorace 2011). The second aim of the study was thus to identify the potential variation existing between monolinguals and bilinguals in subject distribution.

The research was based on elicited data from semi-spontaneous production (narratives) and interpretation of biclausal forward anaphora contexts using adapted methodological tools, which have been successfully employed in previous studies (Hickmann 2003; Mastropavlou, Katsiperi, Fotiadou, Fleva, Peristeri & Tsimpli 2014). It is carried out in light of relevant theoretical accounts, such as the Accessibility Theory (Ariel 1990), the Position of Antecedent Hypothesis (Carminati 2002), the Interface Hypothesis (Sorace 2011) and the Vulnerability Hypothesis (Prada Pérez 2018).

The findings of the crosslinguistic comparison showed that NS in Greek and Spanish seem to work in similar directions in both production and interpretation, while significant differences emerged in the scope of OSP, according to predictions. In particular, the Greek OSP is deictically marked because it is identical in form with the demonstrative pronoun, thereby having a narrower scope compared to its Spanish counterpart. Interestingly, the bilingual performance manifested variation in the scope of NS, while production and interpretation of OSP remained unaffected by

5

language contact, contrary to predictions of the Interface Hypothesis. This is presumably due to the deictic nature of the Greek OSP, which renders it relatively categorical, hence less vulnerable. The study took a lifespan approach in considering adult bilingual speakers of all ages. Crucially, the age factor was found to have an impact on the use of NS in bilinguals as well as on interpretation of NS anaphora, with older speakers following the 'recency of mention' strategy. The findings support the idea that NS are refentially complex and involve external interface conditions, hence they are vulnerable.

1.2 Greeks in Chile

Migration of Greeks in Chile forms part of the great European migration wave to the American continent starting in late 19th century and driven by socioeconomic hardship and political unrest. It reached its peak during the period 1880-1960 with an estimate of 100,000 Greeks relocating in South American countries, including at least forty Chilean cities or towns (Tamis 2006, 2009). In Chile, Greek migration, largely voluntary and legal throughout its history, officially started in 1876 (Zorbas & Nicolaides 2010). The size of Greek migration in Chile was relatively small compared to neighbouring countries such as Argentina (Tamis 2006; Zombolou 2011).

Greek communities were officially established in 1916 in Antofagasta and in 1918 in Santiago (Tamis 2006; Burgos Cacharos 2009). The estimate is that more than 3,500 Greeks immigrated to Chile between 1910 and 1963. During the 50s, an increase in the population of Greek origin/ancestry in the metropolitan area was observed due to new arrivals and relocation of Greek families from the provinces to the capital (Tamis 2006). The Greek population in Chile experienced decline during the Chilean dictatorship (1973-1989), when many settlers definitely left the country and at the same time migration was paused (Tamis 2006, 2009). There has been sporadic migration of Greeks to Chile since the 90s while the Greek Embassy was established in Santiago in 1992. The deepening of the current financial crisis in Greece (since 2009) in conjunction with the concurrent prosperity of the Chilean economy has resulted in a new occasional migration flow, which is still in progress.

6

In Chile, there are eight Greek communities founded across the country, with the Greek community of Antofagasta being the oldest and the one in Santiago being currently the largest (Figure 1.1).

Figure 1.1. Map showing the cities of established Greek Communities across Chile

Presently, the Greek community in Chile is culturally and linguistically fully assimilated into the Chilean society. The demographic statistics of its population are inconclusive. In 2003, less than 1,000 Greeks resided in Chile, according to Tamis (2006), while the total size of the population of Greek ancestry in the country is estimated to be larger. In early 2017, access to registration files of the Greek Embassy in Santiago allowed me to pinpoint 519 living Greeks, holders of a Greek passport and permanent residents in Chile, with most of them residing in Santiago. The above-stated number cannot accurately correspond to actual speakers of Greek. In Chile, Greek passport holders who are not first-generation immigrants do not necessarily speak the language, while there are speakers of Greek descendance without a Greek passport, as well as L2 speakers with no Greek background.[2]

---

[2] On Greek diaspora and its definition see Damanakis (2007) and Chatzidaki (2016).

Greeks who have been living in Chile for many years do not (often/regularly) visit Greece and several of those born in Chile have never been to Greece. The most recent arrivals, involving younger generations of Greeks, maintain stronger bonds with their country of origin, substantially facilitated by the emergence of electronic media and the widespread use of the Internet.

The majority of Chilean Greeks and first-generation Greeks have entered into mixed marriages with Chileans. Children born to Greek parents who are first-generation immigrants speak the language as simultaneous bilinguals, i.e. they are HS. In 2018, in Santiago there were 17 children HS up to 8 years old belonging to 11 binational families known to me. Children born to second- or third-generation of Greeks rarely or scarcely speak the language. A small number of the latter children attend language courses at the Greek community of Santiago taught once a week in the Saturday school. Ethnic pride is generally instilled by families of Greek background into their offspring while Greece also enjoys positive attitudes among Chilean people. It could be said that the Greek language in Chile has a high prestige, but no actual utility and no visibility (see Coulmas 2013, 2018). Religious duties are performed in Spanish in the Greek Orthodox Church located in Santiago (the only one in Chile). Finally, it is worth noting that the present study is conducted in a society where bi- or multilingualism is the exception and not the norm.

1.3 Dissertation structure

The present dissertation is divided into nine chapters, this introduction being the first one. Chapter 2 sets forth the theoretical views taken into consideration in the present study, i.e. the syntactic, referential and pragmatic properties of subjects in NS languages. The chapter also describes the pronominal paradigm of Greek and Chilean Spanish as well as the cases of verbal syncretism in these languages since this is potentially related to the use of subjects.

Chapter 3 focuses on specific theoretical hypotheses proposed in order to explain the bilingual behaviour, such as representational, processing and vulnerability accounts. This part also reviews relevant studies in which overuse of overt subjects or NS has

been attested. Chapter 4 subsequently considers previous research that has been conducted on Greek and Spanish production and interpretation of referential subjects. Findings in the aforementioned studies and relevant theoretical perspectives are taken together in order to formulate the predictions of the present study. Chapter 5 thoroughly describes the methodology followed for the present research, including a description of the tasks administered, the research participants, the data collection process and other methodological considerations.

Chapter 6 presents the production task (Study 1), i.e. the research questions and predictions followed by the results with both quantitative and qualitative analyses of the data and then a critical discussion of the findings. Chapter 7 presents the interpretation task on anaphora resolution (AR) (Study 2) following the same structure. The results from the two studies are taken together and discussed in Chapter 8, which contains the general discussion. In particular, the crosslinguistic variation in Greek and Spanish, as well as the (non-)variation of the adult bilingual systems are discussed engaging with previous related literature. Aspects considered include discursive strategies of reference production, ambiguity resolution, use of defaults, as well as the role of age at the time of testing. Theoretical implications and sociolinguistic considerations are also discussed in Chapter 8. Finally, Chapter 9 offers the main conclusions drawn from the study, its limitations and lines for future work.

2. Theoretical framework

2.1 Approaching language acquisition

The question of how we acquire language is central in the discipline of linguistics and different approaches have been put forward to explain the phenomenon. Universal Grammar (UG) was offered as one possible answer from the formal/generative position. Humans are born with the UG system, a biological linguistic endowment, consisting of two types of constraints: principles and parameters. Principles encode the invariant or universal properties of languages, which make languages similar, while parameters represent the variation dimension encoding properties that vary from one language to another. Parameters are associated with lexical items and particularly with functional categories (UG lexicon). Crosslinguistic variation (parameterisation) is attributed to differences in morphological features of functional heads as specified in the lexicon. L1 acquisition entails parameter setting to the value expressed by the language of the environment. For Biberauer and Roberts (2012), parameters are not provided by UG but are emergent properties resulting from interaction of a minimal UG, experience of primary linguistic data (input) and third factors (Chomsky 2005). Third factors are principles not specific to language faculty, with computational efficiency being the hallmark of their effects relating to processing and general cognitive limitations (e.g. memory) or economy principles (Westergaard 2014; Lohndal & Uriagereka 2016).

The existence of parameters explains the efficiency of L1 acquisition given the great complexity of the human language system. The setting of parameters takes place very early in life and parameters which are most frequently used, such as the NS parameter (NSP), are set before others (Wexler 1998). The general ultimate success in L1 acquisition does not mean that the process is error free, yet developmental syntactic errors found in child language are sparse and gradually eliminated without explicit intervention, unlike adult L2 learning (Montrul 2004b).

Input from sufficient primary linguistic data is indispensable for language acquisition. As formulated by Montrul (2016b: 1), 'input consists of actual samples of language use or naturally occurring written/oral discourse, preferably experienced in the context of social interaction'. Input contains properties which trigger the selection of the correct parameter value (Ayoun 2003; Valian 2015). In order to set a parameter to the right value, children need to be exposed to minimal amount of relevant examples from the environmental input, which may be defective ('poverty of the stimulus'). The role of input is not always clear since research has revealed robust usage patterns in child language that deviate from the input. For instance, children acquiring a NS language (e.g. English) go through a stage in which they omit subjects in finite clauses (see e.g. Hyams & Wexler 1993; Hyams, Mateu, Ortfitelli, Putnam, Rothman & Sánchez 2015; Liceras & Fernández Fuertes 2017). Given that adults providing input do not generally omit subjects, the frequency of NS instances in child speech constitutes evidence against frequency effects postulations (Yang 2015).

As Tsimpli (2014) explains, in monolingual development 'core' grammatical components are viewed as narrowly syntactic phenomena and are acquired earlier than those requiring components located outside narrow syntax, such as semantics, pragmatics and language-external cognitive resources. Accordingly, while NS are mastered and used in appropriate ways since early years (around the age of 2), other linguistic aspects, such as pragmatically conditioned use of pronouns and ambiguity resolution, are acquired later because of their complexity. 'Late' phenomena are associated with discourse properties, which are demanding due to their sensitivity to language-external domains and real-time processing constraints (Tsimpli 2014: 286). The role of input is particularly important for the acquisition of 'late' phenomena.

Thus, inborn language capacity is not enough by itself for language acquisition; an environment providing direct contact with the target language (input) is also required (Yang 2015). Children then parse the input of a specific language selecting relevant values and building syntactic structure based on the principles provided by UG. According to Westergaard (2014), this parsing results in children discovering micro-cues, i.e. small pieces of abstract syntactic structure, which form part of their knowledge of the specific language that they acquire.

Despite the consensus that both biological and social factors are needed for achieving full native language attainment, the relative contribution of 'nature' and 'nurture' is a prominent debate in the field of language acquisition. The UG theory gives prominence to the innate linguistic faculty, based on the fact that input alone cannot account for all the sentences that children produce and understand. On the other hand, emergentist and cognitive approaches hold that language is acquired mainly from interaction with the environment, i.e. by analogy, pragmatics, social interaction and statistical frequencies in the input, in pace with general cognitive development. Linguistic complexity and abstract grammar are shaped by input and emerge by exposure to statistical regularities present in ordinary linguistic data. In the so-called usage-based approach to grammar acquisition, 'pattern-finding' is a central cognitive construct (Tomasello 2015). This 'processor' underlying language acquisition tracks frequency, creates associations, computes distributional contingencies and also minimises the burden on working memory (O'Grady 2008: 458). From this perspective, in order to understand language acquisition, the aim is to capture 'the physical and psychological processes of perception, attention and memory' (Ellis 1998: 656).

The two general approaches to language acquisition are neither mutually exclusive nor irreconcilable. Although their premises are intrinsically different, they are arguably complementary because neither of them alone can puzzle out the complexity of language. Innate principles as well as environmental and cognitive factors are deemed crucial for language acquisition. In words of Yang (2004: 451), 'it is a truism that *both* endowment and learning contribute to language acquisition [...]. Consequently, both must be taken into account, explicitly, in a theory of language acquisition' (emphasis in the original). It is therefore beneficial to seek common ground among the distinct theoretical perspectives. As Sorace (2011: 25) argues, 'linguistic theory is not sufficient by itself to explain the interaction of linguistic and non-linguistic factors that appears to underlie optionality in linguistic behaviour'. Optionality is looked into in relation to subject distribution in Chapter 3.

2.2 On the Null Subject Parameter

Subject distribution has been extensively studied within the principles and parameters model. Principles are general and invariant, hypothetically common to all languages, while parameters determine structural variability among languages. Principles can be general statements, such as the Extended Projection Principle (EPP), according to which every sentence must have a subject. Parameters are options that allow for variation and are typically two-valued: each language sets the parameter to the target value (positive or negative setting). An outline of the NSP is given subsequently in order to approach the nature of NS.

2.2.1 Analyses of null subjects

The observation that some natural languages allow the possibility of leaving a referential subject of a finite verb phonetically empty (unexpressed) was captured by the NSP (Rizzi 1982, 1986) leading to the generalisation of dividing languages into two types: those that require subjects to be phonetically realised (non-NS languages) and those that do not (NS languages). The NSP, as put by Tsimpli et al. (2004: 263), 'is an example of a parameter which, on the one hand, determines syntactic differences between languages, and, on the other, "feeds" interpretable features like topic, focus, and definiteness, which regulate the distribution and interpretation of the formal options'. A recent typology of the NSP (Biberauer, Holmberg, Roberts & Sheehan 2010) suggests four types of NS languages: expletive NS languages (e.g. German), partial NS languages (e.g. Finnish), discourse (radical) pro-drop languages (e.g. Chinese) and consistent NS languages (e.g. Greek, Spanish).

The NSP is derived from the EPP. The canonical subject position where the EPP is satisfied is the Spec TP position within the XP, as shown in (1). Specifier (Spec) is the immediately dominated element by XP and sister to X'. Tense Phrase (TP) stands for the phrase that carries the verbal inflection, i.e. tense and agreement features such as person and number (Agr).

(1)

```
                    CP
                   /\
                    C′
                   /\
                   TP
                  /\
   Spec TP position   T′
                     /\
           [±Tense]    VP
           [±Agr]     /\
                     V    DP
```
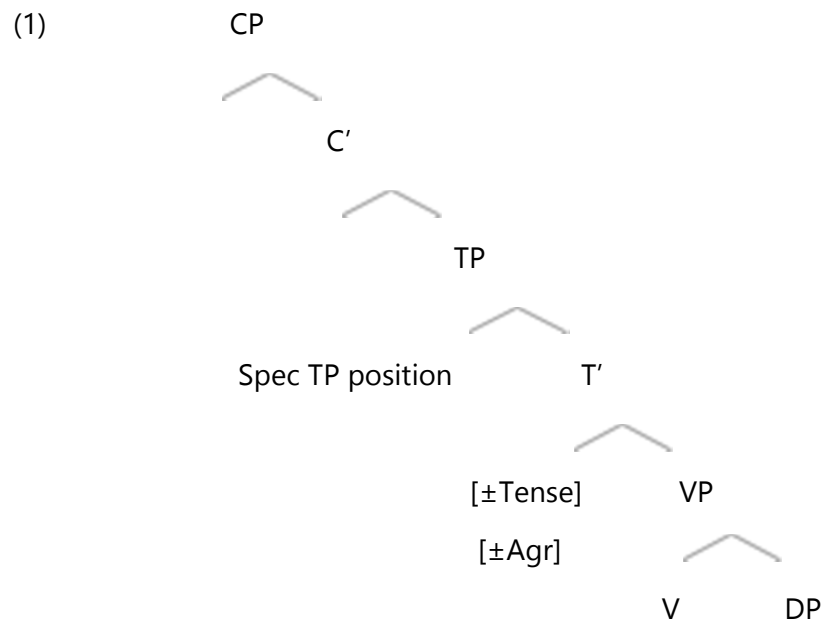
The EPP is a universal principle, i.e. a general grammatical property of sentences, thus it is not specific to particular languages or types of clauses. The EPP is explained as a strong D feature on T that needs to be satisfied by another element with the same D feature (i.e. a nominal category, a noun phrase or DP acting as the subject). Whether the Spec TP position is filled by overt elements depends on the language. The typical paradigm of a non-NS language is English, in which subject pronouns must always be phonetically articulated. In English, an overt element in Spec TP is required and can be either the external argument of the verb (if the verb assigns an external theta role) or an expletive (e.g. *it*, *there*).

Among the different types of NS languages, most Romance languages, including Spanish, as well as Modern Greek are consistent NS languages allowing referential subjects of finite clauses to have overt or null forms in discourse-appropriate contexts (Roberts & Holmberg 2010). Consistent NS languages typically have rich agreement inflection on finite verbs. According to usual conventions, 'rich' means bearing enough morphology to yield non-ambiguous person and number information for the subject (distinct personal endings on the verb) in all tenses. As indicated by Tsimpli et al. (2004: 259), 'agreement features are morphologically realised in a rich paradigm which interacts with Tense distinctions, whereas English has no corresponding richness in the inflectional domain'.

In NS languages, the Spec TP position does not require to be filled by an overt element; however, the EPP requires a non-overt subject in Spec TP. This non-overt subject is known as *pro* and constitutes a phonetically null element present whenever a finite verb has no apparent subject.

There are many proposals to account for the way in which *pro* is licenced (e.g. Borer 1986; Rizzi 1986; Alexiadou & Anagnostopoulou 1998; Sheehan 2006). Setting aside the particularities, the views converge in that in NS languages the inflectional head must be pronoun-like. Roberts and Holmberg (2010: 14) formally encode this property as the presence or absence of a D-feature associated with T. In consistent NS languages, T bears a D feature, whereas in non-NS languages it does not. Whatever the specific syntactic analysis of the NS phenomenon, it has rather little bearing for the immediate purposes of the present study. The major observation taken into account is that syntactic licencing of NS co-exists with OSP in both Greek and Spanish. As Belletti, Bennati and Sorace (2007: 658) note, although the availability of NS 'is traditionally regarded as a direct consequence of the positive setting of the NSP [...], this view now appears to be only partly correct, since discourse factors also play a crucial role'. Discourse factors are explored in later sections in more detail.

2.2.2 Syntactic properties associated with the NSP in Greek and Spanish

Apart from rich agreement morphology, the positive value of the NSP has been traditionally associated with a well-known cluster of derived grammatical properties, such as null expletives, inverted subjects and apparent absence of complementiser-trace effect (see e.g. Camacho 2013). The shared property in Greek and Spanish examined in the present study is the co-occurrence of null and overt subjects.

Both Greek and Spanish allow the suppression of referential subjects and, more specifically, subject pronouns may have overt or null forms. Referential (or thematic) subjects refer to concrete or abstract entities in the world and can have the form of a lexical subject (LS) (lexical noun phrase or full DP) as in (2), a personal pronoun as in (3) or a NS as in (4). Pronominal subjects can be demonstrative pronouns as well and LS can be proper names or demonstrative phrases.

(2)   a. *O mathitis pijeni sto sxolio.*   (Greek)

       b. *El alumno va a la escuela.*   (Spanish)

        'The student goes to school.'

(3)   a. *Aftos pijeni sto sxolio.*   (Greek)

       b. *Él va a la escuela.*   (Spanish)

        'He goes to school.'

(4)   a. ∅ *Pijeni sto sxolio.*   (Greek)

       b. ∅ *Va a la escuela.*   (Spanish)

        '[He/She/It] goes to school.'

In both languages, pronominal subjects can be indefinite (e.g. *kanenas* in Greek; *nadie* in Spanish: 'nobody'), quantifiers (e.g. *oli* in Greek; *todos* in Spanish: 'all') and interrogatives (e.g. *pjos* in Greek; *quién* in Spanish: 'who'), among other possibilities. These, however, are not examined in the present study.

Lexical and overt pronominal subjects can be dropped in both languages because of the rich and uniform verbal inflectional morphology which licenses NS (see Jaeggli & Safir 1989 on Morphological Uniformity Principle). The referent of a NS can be thus identified through the verbal inflection but also through discourse factors when morphology fails (e.g. Belletti et al. 2007; Camacho 2013). As will be explored in more detail, null and overt pronominal subjects are not in free variation, i.e. they are not strictly optional, since their distribution is regulated pragmatically in the discourse.

Greek and Spanish also allow 'free' subject-verb inversion, which is related to discursive factors. The default, unmarked, or canonical word order in unfocused declarative contexts in both languages is SV(O), but orders VS(O) and VOS are also permitted. The position of subject may be preverbal or postverbal, depending on transitivity, verb class, information structure and definiteness (Perlmutter 1978; Tsimpli et al. 2004; Lozano 2006b; Tsimpli 2007).

2.3 Distribution of subjects in Greek and Spanish

Subject distribution has been investigated in Greek and Spanish in monolingual and different types of bilingual speakers and communities from various theoretical perspectives. There are well defined syntactic proposals that have examined the structural properties for the presence and absence of subject forms. Crucially, subject expression and interpretation is also constrained by the interplay between syntax and semantic, pragmatic, cognitive and discourse conditions. Although these conditions apply universally, they are also language specific. This section aims to elucidate relevant aspects of Greek and Spanish subject distribution, starting from defining discourse, reference, deixis and anaphora.

2.3.1 Discourse and reference

Discourse is understood as any coherent succession of clauses. The clause as the basic information-processing unit of discourse is instantiated in utterances, which are organised in sequences by language users to express meaning (e.g. Grosz, Joshi & Weinstein 1995). As a general rule, in order for utterances to be processed and interpreted, recourse to contextual information is needed. Context, according to Ariel's (1990: 5) 'geographic' division, is distinguished by the nature of informational material, which corresponds to three types: 'General or Encyclopaedic Knowledge, the Physical Environment of the speech event, and the Linguistic Context, i.e. previous utterances actually mentioned in the discourse'.

Reference is the relation between a linguistic expression and the concrete object or abstraction that it designates, or the relation between two linguistic elements in which one provides the information necessary to identify or interpret the other (e.g. Kibrik 2011; Sullivan 2012). Zulaica-Hernández and Gelormini-Lezama (2017: 1) define reference as an essential component of the human cognitive system, which concerns 'the human ability to refer to physical, textual and mental entities by means of linguistic expressions, also known as referring expressions'. Thus, such expressions are linguistic elements that 'perform a mention of a referent', which is a concept in language users' mind understood only within a discourse context (Kibrik 2011: 5).

18

Depending on informational load and referential specificity, referential expressions can be divided into lexically full and reduced devices (Kibrik 2011). As regards referential subjects, LS are nominal elements that contain a large amount of information. They are used in discourse to introduce new information (focus), which requires the explicit (as opposed to anaphoric) mention of constituents. LS can be also used to reintroduce old information in cases where explicit disambiguation is needed. Reintroduction and maintenance of information into the discourse is usually made through anaphoric means and specifically by employing overt or null pronouns (zero forms), the latter being the most economic form to establish anaphora to a prominent antecedent. Overt and null pronominal subjects, along with bound personal pronouns (clitics), are reduced referential devices (Kibrik 2011).

Concisely, in NS languages the basic morphosyntactic forms that a subject takes are:

(a) full (definite and indefinite) noun phrases, i.e. lexical subjects (LS)
(b) overt subject pronouns (OSP)
(c) null subject pronouns (NS)

Speakers do not arbitrarily switch between different possible referential forms, but these are used following general patterns determined by contextual factors. A plausible question could be why OSP are ever employed in a language that allows NS, especially taking into account the rich morphology in languages such as Greek and Spanish that generally permits subject identification. The answer to this question will become clear soon.

'A pronoun is a word which takes the place of a noun phrase and can function as the subject, complement or object of a verb, or as the object of a preposition' (Holton et al. 2012: 112; see also Bhat 2004). A pronoun is seen as an element which lacks inherent reference and is co-indexed with an antecedent from which it receives its referent, hence its full meaning. Thus, the distribution of pronouns is intrinsically related to discourse-linking identification by referring to an antecedent; hence, to contextual information.

Pronouns in Greek and Spanish are marked in a three-person distinction (first, second, third), number (singular, plural) and gender (masculine, feminine, neuter). The OSP examined in this study are the personal and demonstrative pronouns in third person. Before describing the pronominal systems of Greek and Spanish, two relevant concepts related to pronouns and their referential properties should be clarified: deixis and anaphora.

2.3.2 Deixis and anaphora

The concept of deixis applies to reference by means of an expression whose interpretation requires the (extra)linguistic context of the utterance in order to convey meaning. It thus concerns the relationship between a linguistic structure, such as a pronoun, and its context of use (Huang 2014, 2017). 'Deixis entails "nongivenness" in the context, which therefore requires (gestural) ostension and precise local specification' (Manolessou 2001: 136). The deictic feature is a distinguishing feature of demonstrative pronouns (Lyons 1999; Alexiadou, Haegeman & Stavrou 2007).

Anaphora involves 'givenness' presupposing presence of the referent in the linguistic context (and less commonly so in the extralinguistic context) (Manolessou 2001: 136). Pronominal anaphora determines the relation between a pronoun and its referent, which is usually present in the same or in an earlier sentence in discourse. Thus, a pronoun with anaphoric function involves sameness of reference (coreference) to a previously mentioned linguistic entity. Less often, it can also anticipate an entity mentioned later (cataphora). The meaning of an anaphoric pronoun thus depends on its antecedent. As Frana (2017: 206) puts it, 'an anaphoric pronoun is probably the prototypical case of a linguistic expression whose full meaning is determined by making reference to a previous referring expression'. Coreferentiality relations between anaphoric pronouns and their antecedents, i.e. anaphoric dependencies, establish and maintain coherence in discourse (Gelormini-Lezama & Almor 2011; Iraola 2015). Anaphora resolution (AR), in particular, is the process of identification of a pronoun referent, i.e. assigning an antecedent to a pronoun (see e.g. Iraola 2015).

2.3.3 Pronominal paradigm of Greek and Spanish

Deixis and anaphora with respect to person are lexicalised in the grammatical category of personal pronouns used for deictic and anaphoric reference. Although subject pronoun distribution is relatively similar in Greek and Spanish, differences worthy of attention can be observed (see details in §2.3.6 and §2.3.7).

The paradigm of personal subject pronouns in Greek and Chilean Spanish have the standard pronominal forms shown in Table 2.1. In both languages, third-person pronouns bear overt gender features. In principle, they may refer to both human and non-human entities, apart from the third-person neuter *ello* in Spanish, which exclusively refers to inanimate entities and has no plural. That said, abstract and inanimate referents are preferably referenced by LS or NS and usually not with OSP. First and second person pronouns do not bear gender features, except the first-person plural *nosotros* and *nosotras* in Spanish, and can only be [+human] realising the speaker and the addressee of a speech event respectively.

Table 2.1. Personal pronoun paradigm in Greek and Chilean Spanish

		Greek		Spanish
Number	**Person**		**Mode**	
Singular	1st	*eγo*		*yo*
	2nd	*esi*	familiar	*tú*
		-	formal	*usted*[3]
		-	informal	*vos*
	3rd	*aftos*		*él*
		afti		*ella*
		afto		*ello*
Plural	1st	*emis*		*nosotros*
				nosotras
	2nd	*esis*		*ustedes*
	3rd	*afti*		*ellos*
		aftes		*ellas*
		afta		-

---

[3] Although the pronouns *usted* and *ustedes* function as third-person pronouns in terms of grammar (they are selected with verbs in third person), they always refer to the interlocutor(s) in the discourse, i.e. to second-person singular and plural respectively (cf. RAE 2018). The singular form *usted* is exclusively formal, while the plural form *ustedes* is used in both familiar and formal speech in Chilean Spanish.

Personal pronouns are so named because they have the grammatical features of person, which are also expressed in verbal inflection and relate to agreement. Personal pronouns can display different properties cross-linguistically but also within the same language (e.g. Carminati 2002; Filiaci 2011; Mayol 2012). Some basic assumptions on the grammatical category of person are introduced next.

2.3.4 The grammatical category of person

Person is an important grammatical feature which determines the use of subject pronouns (e.g. Carminati 2005; Otheguy & Zentella 2012; Peskova 2013; Shin 2016). There is a functional distinction among the different persons of pronouns, which differ according to their content (Pinto 2012; Lubbers Quesada 2015; Schmitz & Scherger 2017; Prada Pérez 2018). The grammatical category of person distinguishes speakers and addressees from each other and from non-participants in the speech event (i.e. individuals/referents which are neither speaker nor addressee), although this distinction is not always straightforward (see Ackema & Neeleman 2017b). The third person is entirely different from the first and second person respecting reference. Relative frequencies of OSP may also differ for every grammatical person. In Spanish, for example, there is evidence that the rate of third-person OSP is low compared to first and second person (see Soares da Silva 2006; Peskova 2013).

First and second person pronouns are indexical or deictic in nature because their referents change depending on who the speaker is. The morphology of the verb ending alone (with some exceptions) and the discourse context involving the speech participants' presence in the conversation determine their reference. Typically, first person signifies the speaker and second person signifies the addressee(s). Denoting and indexing the participants of the speech act, first- and second-person pronouns are in principle [+specific][4] and [+human/animate] and do not stand for an actual noun phrase. Since the referents are normally present in the extralinguistic context, they are highly accessible or contextually salient, thus the structural environment required for their use is minimal (Pinto 2012; see also Steward 1999).

---

[4] Second-person singular pronouns can be also used to express generic reference in both Greek (Holton et al. 2012) and Chilean Spanish (Mayol 2012).

On the other hand, third-person pronouns are anaphoric in nature and refer to lexical or proper noun phrase referents, animate or inanimate, introduced in the preceding (or following) discourse, bearing the same features of the corresponding noun phrases. Although the form of third-person pronouns may be generally analogous to first- and second-person pronouns, third person differs from first and second person in the way of establishing reference, i.e. in referent accessibility (Carminati 2005). Third-person entities are often absent from the extralinguistic context, hence less accessible and more difficult to track (Shin & Cairns 2012). Competing referents may thus exist for third-person pronouns, but usually not for first and second person. Third-person pronouns are regarded as a type of 'discourse tracking devices' (Pinto 2012: 290) or 'reference tracking mechanisms' (Shin & Cairns 2012: 10).

In brief, the basic difference between first/second person and third-person pronouns is based on context: interpretation of the former is inherently connected to the situational information or extralinguistic context (hence deictic), while interpretation of the latter is independent of it (hence anaphoric) (Lozano 2009; Pinto 2012; Lubbers Quesada 2015; Schmitz & Scherger 2017; Prada Pérez 2018). The relatively low frequency of third-person pronouns is attributed to the fact that these are anaphoric, i.e. interpreted as given information (Peskova 2013). Reference tracking is crucial for third-person referents as compared to that for first- and second-person because in the former case there is a risk for ambiguous reference, whereas in the latter case reference is virtually always transparent.

2.3.5 Acquisition of pronouns

Pinto (2012) explains that third-person pronouns are noun phrases which become DPs (referring expressions) when combined with the grammatical feature of definiteness. This is why third-person pronouns have a more articulated internal structure than first- and second-person pronouns, which do not contain such descriptive features. The different nature of pronouns' features results in an asymmetrical acquisition schedule of person forms in monolingual children acquiring NS languages. Children acquire first and second person singular before the third person singular in Italian (Pinto 2012). The same pattern emerges in Spanish where

overt third-person subject pronouns are not found in very young children's speech (Bel 2003; Shin & Cairns 2012; Shin 2016). Aside from definiteness, other fundamental grammatical notions which must be in place before third-person pronouns emerge are phi-features distinctions and agreement (Pinto 2012; see also Hickmann 2003). A similar asymmetry is found in Spanish L2 learners in the production of first and second person, which is similar to monolingual native speakers, whereas production of third-person animate subject pronouns appears later and seems to be more vulnerable in language contact situations (Lozano 2009). In what follows, singular third-person subject pronoun of Greek and Spanish will be presented in more detail.

2.3.6 Third-person overt subject pronouns in Greek and Spanish

2.3.6.1 Greek

In Greek, subject personal pronouns have strong and weak forms. NS are classified as weak pronouns and constitute the default form. The strong forms are the 'marked' option and function as the subject of a clause 'when emphasis or distinction is required' (Holton et al. 2012: 113). The strong form of the third-person personal pronoun is *aftos* and is inflected in masculine, feminine and neuter (grammatical) gender. These forms are also inflected for number and case (nominative, accusative and dative which surfaces as genitive), the nominative case indicating the subject of a verb. Third-person pronouns in Greek are in essence forms of the demonstrative *aftos*[5] (e.g. Manolessou 2001; Tsimpli & Sorace 2006; Mavrogiorgos 2010; Tsimpli 2011; Holton et al. 2012; Prentza & Tsimpli 2012).

Table 2.2. Third-person singular pronoun *aftos* in Greek

Greek third-person singular pronoun *aftos* in nominative case			
Gender	Masculine	Feminine	Neuter
Personal pronoun (strong forms)	*aftos*	*afti*	*afto*
Demonstrative pronoun	*aftos*	*afti*	*afto*

---

[5] Through a diachronic lens, it can be seen that initially *aftos* was used only anaphorically. Due to a pragmatic shift from anaphoric to deictic functions, it eventually acquired the uses of a fully-fledged demonstrative pronoun (exophoric, cataphoric), which obtains in Greek today (Manolessou 2001).

Apart from *aftos*, there are two more demonstrative pronouns in Greek: *ekinos* and *(e)tutos*, the latter being used similarly to *aftos* in its deictic usage yet much less often. The demonstratives may operate as:

(a)   strong third-person pronouns, devoid of deictic content, e.g. *aftos majirepse* 'he cooked'; in order to keep their anaphoric function, personal pronouns are unstressed; if a personal pronoun carries prosodic stress, it then has a demonstrative function with further deictic properties (Tsimpli et al. 2004), e.g. *AFTOS majirepse* 'HE cooked'

(b)   demonstrative pronouns, with deictic properties, e.g. *aftos ine o mathitis* 'this is the student'

(c)   demonstrative determiners in adjectival position modifying a noun, thereby forming demonstrative phrases, e.g. *aftos o mathitis* 'this student'; these cases fall into the category of LS

With respect to (c), i.e. when used as determiners/adjectives, Manolessou (2001) argues that *aftos* and *ekinos* express the distinction of deictic vs anaphoric through alternative word orders: deixis involves the demonstrative in the pre-article position (e.g. *aftos/ekinos o mathitis*) and anaphora placement in the post-nominal position (e.g. *o mathitis aftos/ekinos*). In short, demonstratives can be used either as pronouns or as determiners within a noun phrase with a deictic or an anaphoric usage in both cases (Mavrogiorgos 2010). Since (c) involves noun phrases (LS), in what follows the focus is on (a) and (b) concerning pronouns.

Demonstratives inherently link to the notion of proximity and distance in the space/time or metaphorical perception within the discourse in relation to the speaker, who is said to constitute the reference point, or deictic centre. As regards the two most commonly used demonstrative pronouns in Greek, *aftos* entails nearness hence proximal, while *ekinos* entails distance hence distal. Dimitriadis (1996: 10) maintains that *ekinos* 'is frequently used with antecedents that did not occur in the previous sentence at all'. Having this proximal-distal, or recent-remote contrast, *aftos* is more commonly used than *ekinos* as third-person OSP, while both are surpassed in relative frequency by NS (Dimitriadis 1996).

As previously shown, Greek has the lexical distinction of proximity vs distance and the syntactic distinctions of deictic vs anaphoric in its pronominal system. There is no specific form for purely anaphoric function, since *aftos* is not used in exclusively anaphoric contexts. Rather, the deictically marked form (the demonstrative) occurs anaphorically. In words of Tsimpli (2011: 102), 'the overt subject pronoun in Greek [...] may have anaphoric uses with a non-topical antecedent (when unstressed), but also deictic or emphatic uses when stressed, in which cases it can co-refer with the subject topic of the sentence'. It can thus denote topic continuity but only when discourse-related features, such as contrast or emphasis, are involved (Prentza & Tsimpli 2012).

In addition, presumably by virtue of its strong deictic component, the use of *aftos* in certain contexts may bear a pejorative nuance when referring to an actual person (e.g. *Irthe aftos* 'He arrived'). Due to this connotation, speakers are careful in using the pronoun *aftos* in order to avoid the impolite overtone which it may transmit. In brief, omission of third-person subject pronouns is encouraged as a form of politeness. The last-mentioned pragmatic effect together with the inherent deictic nature of third-person pronouns and the distinction of person and number in verb morphology may shed some light into why OSP particularly in third-person singular are 'relatively rare' in Greek (Dimitriadis 1996: 7; see also Chiou 2012).

2.3.6.2 Spanish

Spanish has stressed pronouns and NS, the former being inflected for case (see Luján 1999b; Ordóñez 2012). The third-person form of the personal pronoun is used as the subject in nominative case and can be of masculine, feminine or neuter gender. In nominative case, the forms *él*, *ella*, *ello* are inflected for number. There is no plural for the neuter form (*ello*) and it can only refer to inanimate entities. NS in Spanish are also viewed as the default forms, as in Greek, while OSP are the marked options (Liceras et al. 2010; Tsimpli 2011; Liceras & Fernández Fuertes 2017).

Since in Greek the demonstrative is important in order to understand the nature of subject pronouns, it is appropriate to consider it in Spanish as well. As opposed to Greek, in which the personal and the demonstrative pronoun *aftos* coincide in form,

in Spanish the demonstrative pronouns are distinctly different from the personal pronouns. There are three inflected demonstratives that can be used most frequently in subject position: *este*, *ese* and *aquel*, which mark gender and number as the personal pronouns do. Neuter forms (*esto*, *eso*, *aquello*) are pronouns referring to inanimate/abstract entities and lack plural. The status of masculine and feminine demonstratives may change: when accompanying a noun (e.g. *esta idea* 'this idea') the demonstratives are determiners; when they stand alone they may function as a pronoun[6] or a determiner. The demonstrative pronoun *este* is the equivalent to the demonstrative *aftos* in Greek.

Table 2.3. Third-person singular pronouns in Spanish

Spanish third-person singular pronouns in nominative case			
Gender	Masculine	Feminine	Neuter
Personal pronoun	*él*	*ella*	*ello*
Demonstrative pronoun	*este*	*esta*	*esto*

In sum, analogously to Greek, the personal pronouns in Spanish may operate as:

(a)   pronouns devoid of deictic content, e.g. *él cocinó* 'he cooked'; if a personal pronoun carries prosodic stress, it then conveys emphasis or contrast, e.g. *ÉL cocinó* 'HE cooked'

The demonstratives carrying deictic properties are used as:

(b)   demonstrative pronouns, e.g. *éste es el alumno* 'this is the student'

(c)   demonstrative determiners in adjectival position modifying a noun, thereby forming demonstrative phrases, e.g. *este alumno* 'this student' (LS)

The demonstratives *este*, *ese* and *aquel* allow a ternary division related to the (physical or metaphorical) distance that separates the speaker from the referent. Traditionally, *este* denotes proximity to the speaker (hence proximal), *ese* denotes proximity to the listener (hence medial) and *aquel* expresses remoteness from both (hence distal). Some analyses postulate that the demonstrative *ese* would be the unmarked element that can take both values of nearness and remoteness used in situations where the relation of proximity is not relevant. In fact, the distance

---

[6] When the masculine and feminine forms function as pronouns, it is usual (though not obligatory) to mark them with a written accent.

established through the lexical distinctions marked by demonstratives is a subjective rather than a real concept, i.e. the expressed distance is always relative and never fixed. Both *ese* and *aquel* (unlike *este*) are also used with a demonstrative value manifesting the so-called deixis *in absentia*, which allows to indicate entities that are neither present in the context nor mentioned in the preceding discourse (RAE 2011). In this regard, *ese* and *aquel* can be seen as equivalents to *ekinos* in Greek, while the demonstrative *aftos* parallels *este* referring to an entity introduced in the immediate (extra)linguistic context.

In Spanish, as in Greek, NS are much more widely used than OSP (e.g. Shin & Cairns 2012). The rate of OSP use is considerably variable in the different geographic varieties of the language (e.g. Lipski 1994, 2012; Otheguy, Zentella & Livert 2007). Caribbean Spanish (Dominican, Puerto Rican and Cuban) displays a high rate of OSP compared to other varieties of Latin American and Peninsular Spanish (e.g. Montrul 2004b; Mayol 2012). In the case of Chilean Spanish, the scarcity of related research does not allow to draw any solid conclusions on the relative frequencies of subject pronoun use. It seems to be the case that third-person subject pronouns in singular and plural are generally used more often than in Peninsular Spanish, but not as much as in Caribbean Spanish, i.e. Chilean Spanish lays in-between the two extremes of the spectrum (see Enríquez 1986; Van Esbroeck 2014).

2.3.6.3 Observations

Drawing on the above and considering Table 2.4 displaying the pronominal forms in nominative singular, the observations on third-person pronoun are the following:

(a)   Since in Greek the personal pronoun is identical in form with the demonstrative, its use may be ambiguous between deictic and pronominal readings; as a result, it has more deictic properties than the Spanish personal pronoun and its use is presumably more restricted than in Spanish.

(b)   In Spanish, there is a three-term distance-related system of demonstrative pronominal forms with the existence of the medial demonstrative (*ese*), which describes an intermediate space between proximal and distal, whereas Greek distinguishes only between the two basic zones (near/far).

(c)　　The neuter forms in Greek are used either as pronouns or as determiners, whereas in Spanish these forms (existing only in singular) are exclusively used as pronouns referring to inanimate or abstract referents, i.e. they cannot modify a noun; this is because in Spanish, nouns are either masculine or feminine, never neuter, and the pronominal use of neuter has very specific referential properties (RAE 2016).

Table 2.4. Third-person personal and demonstrative pronouns in Greek and Spanish

Personal		Demonstrative		
Language		**Proximal**	**Medial**	**Distal**
Greek	*aftos*	*aftos*		*ekinos*
	afti	*afti*		*ekini*
	afto	*afto*		*ekino*
Spanish	*él*	*este*	*ese*	*aquel*
	ella	*esta*	*esa*	*aquella*
	ello	*esto*	*eso*	*aquello*

Overall, there is no strict one-to-one correspondence between Greek and Spanish third-person overt pronominal subjects. However, in both languages OSP are said to always convey a semantic or pragmatic value (e.g. Haegeman 1994; Chiou 2012; Liceras & Fernández Fuertes 2017; see also §2.4 and Chapter 4).

2.3.7 Ambiguous person morphology in Greek and Spanish verb forms

2.3.7.1 Greek

Verb forms with ambiguous person morphology are found in Greek, albeit to a lesser extent compared to Spanish, potentially hindering the identification of a NS referent.

(a)　　The verb *ime* 'to be/exist' is ambiguous between third person singular and third person plural in all verb tenses and moods as exemplified below:

aftos/afti/afto / afti/aftes/afta　　　*ine* 'is/are'
'he/she/it / they'　　　*itan(e)* 'was/were'
　　　　　　　　　　tha ine 'will be'
　　　　　　　　　　tha itan 'would be'
　　　　　　　　　　na ine '(to) be'
　　　　　　　　　　na itan '(to) have been'

(b)   Second-conjugation verbs (oxytone type B; see Holton et al. 2012) in passive imperfect are also ambiguous between third person singular and third person plural, e.g. *aftos theorundan(e) / afti theorundan(e)* 'he was considered / they were considered', and likewise with the conditional, e.g. *aftos tha theorundan(e) / afti tha theorundan(e)* 'he would be considered / they would be considered'.

The frequency of verb forms with morphological overlap in (b) is very low in Greek, unlike the forms of *ime* ('to be/exist'), as seen in (a), which are highly frequent.

2.3.7.2 Spanish

It has been observed that there may be restrictions on the availability of NS in Spanish by virtue of the inability of certain verbal forms to identify their subject referent due to verbal syncretism (Silva-Corvalán 1994; Zagona 2002; Filiaci 2011; Camacho 2013; Shin 2014; Duarte & Soares da Silva 2016; see also Cardinaletti 2014).

More specifically, the morphology of verb desinences in Spanish is neutralised and is identical, hence potentially ambiguous, in first and third person singular, as well as formal second-person singular, in less distinctive verb paradigms such as imperfect, pluperfect, conditional and all subjunctive forms (Table 2.5).

Table 2.5. Example of morphological neutralization of Spanish verb forms

TAM	*yo / él / usted*	'I / he / you'
past imperfect	*estudiaba*	'was/were studying'
pluperfect	*había estudiado*	'had studied'
simple conditional	*estudiaría*	'would study'
perfect conditional	*habría estudiado*	'would have studied'
present subjunctive	*estudie*	'study/ies'
imperfect subjunctive	*estudiara/estudiase*	'would study'
perfect subjunctive	*haya estudiado*	'would study'
pluperfect subjunctive	*hubiera/hubiese estudiado*	'would have studied'

In Latin American Spanish, verb morphology may be also ambiguous in all tenses and moods between the second-person plural *ustedes* (used to express both formal and familiar 'you') and the third-person plural *ellos/ellas*. For example, *ustedes estudian/estudiarían* 'you study/would study' and *ellos estudian/estudiarían* 'they study/would study'. This is because *ustedes*, as well as *usted*, grammatically function as third-person pronouns due to etymological facts[7]; hence, these pronouns are conjugated with third-person verb forms. However, of all the personal subject pronouns, it is *usted* the one which most frequently accompanies the verbal form due to its formality (Steward 1999; Peskova 2013). Therefore, the question of potential ambiguity mostly lies in the distinction between first and third person singular.

Since Spanish is a NS language, in all the above-mentioned cases the optional absence of overt subject could cause person ambiguity since verb inflection alone does not clearly indicate the intended person. It has been suggested that Spanish speakers tend to use OSP with ambiguous person morphology verb forms (Lubbers Quesada & Blackwell 2009; Shin 2014, 2016; Shin & Erker 2015; Prada Pérez 2018; cf. Filiaci 2011; Van Esbroeck 2014). This is illustrated in (5) from Silva-Corvalán and Enrique-Arias (2017: 173), in which *yo* 'I' is used for reference disambiguation since the verb form *estaba* 'was' is morphologically ambiguous between first and third person singular.

(5) *[...] ella iba a mi lado y **yo** estaba temblando.*
     '[...] she used to go by my side and **I** was shaking.'

Nonetheless, in Filiaci's (2011) study on anaphora resolution (AR) focusing on monolingual Peninsular Spanish, there was no evidence of increased preference for OSP due to ambiguous verbal morphology in the patterns found in anaphora interpretation (see §4.2.1.1).

Additionally, Chilean Spanish is one of the varieties which are characterised by the lenition or deletion of word-final /s/, which is reduced to an aspiration [h] or completely lost (see Lipski 1994, 2012). The weakening or loss of the final /s/ on

---

[7] These forms evolved through phonetic contraction from *vuestra(s) merced(s)* 'your grace(s)'.

second-person singular verb forms may generate (additional) ambiguity between second and third person in present indicative, as well as among first, second and third person in past imperfect, pluperfect, conditional and subjunctive. Native speakers of Chilean Spanish might tend to use more overt subjects to compensate for information loss due to the weakening or deletion of the second-person singular final /s/ in ambiguous verb forms (see Steward 1999). This is not the case in informal registers where the use of *voseo* is widespread, obtaining different second-person singular endings from those of standard Spanish (e.g. *estás* vs *estai* 'you are').

2.3.7.3 Subjunctive

Spanish has indicative, infinitival and subjunctive complements in complementary distribution. The subject pronoun in the indicative complement in (6) is referentially free (Luján 1999a: 105). The subjunctive verb forms (Table 2.5) in complement clauses indicate disjoint reference between matrix and embedded subject (contra-indexation or Obviation), as in (7) (Goodluck, Terzi & Chocano Díaz 2001: 157). If the verb is infinitival, the embedded subject (also known as PRO) expresses same reference (co-indexation or Control), as in (8), i.e. it is obligatorily controlled by the matrix subject (Goodluck et al. 2001: 157; see also Parodi & Tsimpli 2005).

(6)   *Juan$_i$ dice que (él)$_{i/k}$ viene.*
        'John says that he comes.'

(7)   *María$_i$ intenta/quiere que $\varnothing_{k/*i}$ cante.*
        'Maria tries/wants (for) someone else to sing.'

(8)   *María$_i$ intenta/quiere $\varnothing_{i/*k}$ cantar.*
        'Maria tries/wants to sing.'

In Greek, the distribution of embedded subjects is different. A crucial difference is the lack of infinitival forms, hence complement clauses involve either an indicative or a subjunctive verb form (Spyropoulos 2007). Contrary to Spanish, the presence of an overt pronoun in indicative complements signals disjoint reference in non-focused contexts, as in (9). Subjunctive complement clauses are introduced by the subjunctive particle *na* and the verb is fully inflected for agreement with the subject, hence it may

be regarded as finite (Goodluck et al. 2001; Spyropoulos 2007; cf. Iatridou1993). In Greek, a distinction is made based on the lexical semantics of the matrix verb, shown in (10) and (11) (Goodluck et al. 2001: 156). Moreover, in Greek the object of a matrix clause can be also the (semantic) subject of the embedded clause, as shown in (12), which involves an Object Control structure (see Beys 2009: 109).

(9)   *O Janis$_i$ lei oti $\varnothing_i$/aftos$_k$ erxete.*
       'John says that he comes.'

(10)   *I Maria$_i$ prospathi $\varnothing_{i/*k}$ na traɣuðisi.*
        'Maria tries to sing.'

(11)   *I Maria$_i$ theli $\varnothing_{i/k}$ na traɣuðisi.*
        'Maria wants to sing' or 'Maria wants someone else to sing.'

(12)   *Epise$_i$ ton Jani$_k$ na fiji$_k$.*
       'S/he persuaded John to leave.'

In sum, in indicative complement clauses the presence of an OSP does not necessarily indicate disjoint reference in Spanish, but it does so in Greek. Furthermore, while in Spanish the referential distinction is determined by the form of the embedded verb, i.e. whether it is infinitival or subjunctive, Greek distinguishes (non-)coreference based on the semantic class of matrix verb (Goodluck et al. 2001). Thus, in Spanish overt subjects are not (invariably) required to mark change of subject referent in cases of subjunctive complement clauses since coreference with the matrix subject is disallowed (Obviation) (see Luján 1999a; Sánchez-Naranjo 2013). In Greek, on the other hand, the use of subjunctive may indicate both coreference and non-coreference between the matrix and the embedded subject depending on lexical semantics, with subject and object control verbs, such as in (10) and (12) respectively (see details in Iatridou1993; Goodluck et al. 2001; Parodi & Tsimpli 2005; Beys 2009).

2.3.7.4 Observations

Compared to Spanish, Greek presents fewer cases of verbal syncretism and only in third-person singular and third-person plural. In Spanish, inflectional ambiguity appears in more verb paradigms and potentially in all three persons. Thus, there are differences in the relative amount of homophony between verb forms in the paradigms of the two languages[8]. It could be said that Greek has richer inflectional morphology on verbs than Spanish or that Spanish has weak agreement in person-verb paradigm compared to Greek. Spanish conflated verb forms, i.e. with ambiguous person morphology, may favour subject pronoun expression for effects of person disambiguation, especially in contexts involving alternation of first and third person singular. In Greek, the syncretism in the third-person singular and plural of the verb *ime* ('to be/exist') may cause ambiguity on subject referent(s) of these verb forms. However, there is no evidence so far on whether overt subjects in Greek are used (more) with syncretised verb forms to avoid potential ambiguity.

In addition, there is a difference between Spanish and Greek in the use of OSP in indicative complement clauses. The use of OSP in Spanish generally allows both coreferential and non-coreferential intepretations, while in Greek the OSP marks non-coreference in non-focused contexts. Although subjunctive in Spanish may be morphologically ambiguous between first and third person singular, when it appears in complement clauses it marks non-coreference between the matrix and the embedded subject. This is not the case in Greek, in which subjunctive in complement clauses may allow both subject coreference and non-coreference with the matrix subject depending on the lexical semantics of the matrix verb.

---

[8] Third-person singular indicative and second-person imperative verb forms are also conflated in Greek and Spanish in some cases, e.g. *Aftos kitakse / (Esi) kitakse* 'He looked / (You) look'; *él mira / (tú) mira* 'he looks / (you) look'. However, the use of imperative is inherently related to interactive speech acts between a speaker and (an) addressee(s). Thus, in spoken language the communicative situation would hardly allow person ambiguity in using imperative verb forms.

2.4 Referential properties of null and overt subjects

2.4.1 Information structure

Speakers' reference to entities in the world relates to information structure, i.e. the structure of a sentence as a means of communicating information to an addressee (Oxford Reference 2017). Information structure involves both introducing new referents and re-introducing already-mentioned referents into discourse. It guides speakers' selections of nouns, pronouns and other referring expressions following complex discursive rules (Arnold, Kaiser, Kahn & Kim 2013). The most common referring expressions are the nominal and pronominal elements. These elements organise reference in discourse involving further specific linguistic choices, such as choice of definiteness in noun phrases (definite or indefinite) and pronoun use (null or overt, personal or demonstrative).

Choices of referential expressions reflect a distinction between contextually given versus new information. Given information implies information which is familiar, i.e. supplied by the speaker, previously known, assumed or inferred, hence somehow recoverable by the addressee. New information is information that is introduced into the discourse for the first time: it is non-derivable, hence not known to, or assumed by the addressee. Apart from previously-established information in the discourse, relevant world knowledge and inferences are crucial in modulating the effects of information structure on reference (see Ariel 1990).

Referential choices also reflect the distinction between the discursive features of topic and focus. Topic involves previously-established information and 'topicality' denotes referential importance (Givón 2016: 6). For Kiss (1995: 7), the '(discourse-)semantic' concept of topic is identical to the notional subject and often (yet not always) coincides with the grammatical subject, while in NS languages topics are also associated with resumptive pronouns or clitics. Focus is new information, highlighted or non-presupposed. The '(discourse-)semantic' function of focus is twofold: (a) it can denote the part of the sentence which carries new information (wide focus); and (b) it can entail an operator expressing identification (narrow focus) (Kiss 1995: 15).

The information status of a referent in the context of a discourse segment is established by the above-mentioned distinctions of given vs new information, or topic vs focus functions (see also Arnold et al. 2013). Although there are different discourse features which regulate subjects, the present study pays particular attention to the presence or absence of topic shift in discourse. Since topics tend to appear in subject position (e.g. Miltsakaki 2011), it is assumed here that topics are the subjects of clauses.

Definiteness largely concerns noun phrases, i.e. lexical subjects or objects in the framework for this study, and deals with the identifiability and non-identifiability of referents in a given context for the speaker or the addressee. Definite noun phrases are referential phrases in which the existence and uniqueness of referents is presupposed or inferable, while indefinite noun phrases are also referential but they presuppose that the entity presented in discourse is new/unfamiliar (see e.g. Abbott 2006; Belletti et al. 2007; Chiriacescu 2014; Mastropavlou et al. 2014). Definite descriptions and pronouns are not equivalent as to their effect on coherence because they engender different inferences on the part of the addressee (Grosz et al. 1995).

With respect to pronominal reference, use of pronouns enables speakers to refer to established (usually previously mentioned) referents avoiding the need to expressly restate them, thus preventing excessive redundancy or repetition (e.g. Keating, VanPatten & Jegerski 2011; Arnold et al. 2013; Iraola 2015). As stated by Gundel, Hedberg and Zacharski (1993: 274-275), 'different determiners and pronominal forms conventionally signal different cognitive statuses (information about location in memory and attention state), thereby enabling the addressee to restrict the set of possible referents'. Thus, pronouns used anaphorically constitute cues guiding the listener to identify an entity in discourse by providing an amount of information about its cognitive status regulated by the speaker (see Almor 1999). In line with Ariel's (1990) 'geographic' view on context, pronouns typically refer to the Linguistic Context, demonstratives to the Physical Context and proper names to General Knowledge (Clark & Marshall 1981 as cited in Ariel 1990).

In generativist terms, the basic assumption on the overt/null subject use is captured by the Avoid Pronoun Principle, a universal of economy formulated by Chomsky in 1981, which outlines the condition of their distribution. OSP use is restricted to cases where their explicit realisation aims at ensuring recoverability in the discourse, emphasis or contrast. In other words, OSP are produced whenever NS are impossible. NS are used for economy purposes on condition that they can be licensed and identified, since omitting a subject is less effortful than expressing it overtly. The Avoid Pronoun Principle holds generally but it applies only to NS languages. However, its instantiation is language specific. The standard use of OSP in Greek and in most Spanish varieties is found in contexts of switch of subject referent or focus.

Thus, referential pronouns in NS languages are governed by both morphosyntactic and discourse/pragmatic properties, i.e. they are both syntactically and pragmatically constrained. Greek and Spanish pronoun properties are more complex than their counterparts in non-NS languages, such as English, which has a less complex pronoun distribution, not conditioned by discourse factors. Accordingly, OSP (both stressed and unstressed) in NS languages would correspond to stressed English pronouns, while NS generally correspond to English unstressed pronouns, albeit without patterning in exactly the same way (Frascarelli 2007; Miltsakaki 2007; Liceras & Fernández Fuertes 2017).

2.4.2 Topic continuity, topic shift and topic ambiguity

Pronouns can be a source of ambiguity. This is illustrated in the English example in (13), where the subject pronoun (*he*) may equally refer to the subject (*John*) or the object (*Charles*) of the preceding clause, rendering its referent unclear (example from Keating et al. 2011: 197).

(13)   *John saw Charles when **he** was walking on the beach.*

In English, OSP are obligatory in all contexts. Greek and Spanish, as NS languages, have more options for expressing subjects. The aforementioned ambiguous sentence could be given in principle as (14) in Spanish (Keating et al. 2011: 197) and (15) in Greek. Both NS and OSP are permitted options.

(14)   *Juan vio a Carlos cuando ∅/**él** caminaba en la playa.*

(15)   *O Janis iðe ton Kosta otan ∅/**aftos** perpatuse stin paralia.*

The apparently free alternation of overt and null subjects is, however, constrained by discursive factors. Traditionally, the use of NS implies reference to an antecedent which is clearly identified by the context, i.e. given. Thus, NS have been characterised as being the default in NS languages and having simple informational structure. In other words, omission of subjects is the unmarked usage of subject pronouns. The presence of OSP, on the other hand, must be justified. Typically, OSP are used in order to mark change of topic and/or to convey focus (contrast or emphasis) for purposes of clarity or ambiguity avoidance. The overt forms are thus considered to be the marked option and to have more complex informational structure (see e.g. Paredes Silva 1993; Tsimpli et al. 2004; Lozano 2009; Sorace et al. 2009; Tsimpli 2011; Kaltsa et al. 2015; Papadopoulou, Peristeri, Plemenou, Marinis & Tsimpli 2015).

The existence of the same index on the pronoun and the subject of the matrix clause indicates coreference. In example (16) in Greek from Tsimpli et al. (2004: 260), the use of the NS is coreferential with the matrix subject; a NS thus implies a non-shifted interpretation for the embedded subject. In contrast, if the OSP *aftos* is used, the antecedent is not the matrix subject but (in this context) the matrix object.

(16)   *O Janis$_i$ prosvale ton Petro$_k$ otan ∅$_i$/aftos$_k$ ton plisiase.*

         'Janis$_i$ insulted Petro$_k$ when he$_{i/k}$ approached him.'

Consequently, a distinction is established between two basic types of contexts, in which distribution of NS and OSP is regulated for reference to given entities: topic continuity and topic shift (Dimitriadis 1996; Argyri & Sorace 2007; Sifaki & Sitaridou 2007; Lozano 2009, 2016; Sorace et al. 2009; Blackwell & Lubbers Quesada 2012; Shin & Cairns 2012; Mayol 2012; Andreou 2015; Bel & García-Alcaraz 2015; Kaltsa et al. 2015; Montrul & Sánchez-Walker 2015; Papadopoulou et al. 2015; Montrul 2016a; Clements & Domínguez 2016; Georgopoulos 2017, a.o.).

Topic continuity (TC), or subject maintenance in the context of this study, is usually encoded via a NS in both Greek and Spanish. Givón (1983: 55) refers to TC as topic 'predictability', which is 'the unmarked case in human discourse' and, in terms of the psychology of discourse processing, it is easier to process: 'the topic is obviously "more accessible" or "more recoverable" psychologically if it remains the same as in the preceding clause(s)'. Referential continuity entails topic predictability/accessibility.

Topic shift (TS), i.e. subject discontinuity, involves a change of subject referent, typically realised by means of an overt form in both Greek and Spanish. Overt subjects thus function as markers of TS. Givón (1983: 55) views TS as topic 'surprise' (as contrasted with 'predictability'), which 'is the marked case in discourse' and makes the subject/topic 'less accessible' or 'harder to recover' in processing. TS is also referred to as switch reference, signalling disjoint reference (non-coreference or referential discontinuity) of a subject with the subject of the preceding (adjacent) clause (Sifaki & Sitaridou 2007; Gijn 2016; Givón 2016).

Prada Pérez (2009) emphasises the difference between coreferentiality and topic continuation: the former determines whether a subject is coreferential with the preceding subject, thereby taking into account the referent's syntactic function, while this is not the case for the latter. The dimensions of referential coherence considered in the present study are TC and TS. As already stated, topics are assumed to be subjects of clauses, thus TC corresponds to subject continuity (coreferentiality).

Givón (1983: 55) also puts forward the concept of topic ambiguity, which concerns the dimensions of continuity and accessibility: 'the more potential topics in the immediate discourse environment that lay claim to the coreference interpretation of a particular predicate clause, the less accessible or recoverable is the topic in psychological terms'. Givón's view is that difficulty in topic assignment may stem from topic discontinuity (i.e. TS), which creates at least some potential for topic ambiguity, especially when there are two or more potential competitors. Paredes Silva (1993) likewise observes that discourse connectedness plays a role in ambiguity since a shift of topic weakens previous expectations.

Whenever a previously mentioned subject reappears in discourse, reference to it can be made via a LS, an OSP or a NS. In contexts of TC, a NS is considered pragmatically appropriate (i.e. felicitous) due to economy reasons, while the other options are generally considered infelicitous although also grammatical. In contexts of TS, a LS or an OSP are regarded as the pragmatically felicitous options (e.g. Lozano 2009). NS can also be grammatical and felicitous in TS when the likelihood of ambiguity occurrence is none or minimal (Lubbers Quesada & Blackwell 2009; Blackwell & Lubbers Quesada 2012). Apparently, there exists a division of labour between NS and OSP, which contributes to felicity in subjects' use determining a degree of coherence, necessary for unobstructed flowing of communication and processing. However, this division of labour is not always straightforward as it depends on different factors.

2.4.3 The Position of Antecedent Hypothesis

The division of labour of NS and OSP in intra-sentential anaphora was captured by Carminati (2002) in her proposal of the Position of Antecedent Hypothesis (PAH) for Italian, which has been influential since its inception in the study of NS languages. The formulation by Carminati (2002: 57) is the following:

> The Position of Antecedent Hypothesis for the Italian null and overt pronouns in intra-sentential anaphora: the null pronoun prefers an antecedent which is in the Spec IP position, while the overt pronoun prefers an antecedent which is not in the Spec IP position.

The asymmetry between NS and OSP use is prompted by the syntactic position of the antecedent, which in turn determines its prominence and consequent anaphoric resolution, i.e. resolving what/who a pronoun refers to. A preverbal subject in the canonical (Spec IP/TP) position of a sentence is assumed to constitute a referent which is interpreted as the default topic of that sentence. The preverbal subject is considered to be more prominent than other antecedents found in positions lower in the hierarchical syntactic structure, such as object position. The prediction of PAH is that NS are linked more often than OSP to subject position, while OSP are preferentially linked to non-subject (non-topic) antecedents.

The antecedent preference for OSP was, however, less established and more dependent on context than that of NS in Italian, as shown in Carminati's findings. 'Violations' of antecedent bias for OSP (i.e. OSP referring to preceding subjects) were less costly for participants to process than 'violations' of antecedent bias for NS (i.e. NS referring to preceding objects), as in (17) (Carminati 2002: 67).

(17)   *Quando Vanessa$_i$ ha visitato Giovanna$_k$ in ospedale, lei$_{i(/k)}$/$\varnothing_{i(/k)}$ le ha portato un mazzo di rose.*

'When Vanessa$_i$ visited Giovanna$_k$ in the hospital, (she) took her a bunch of roses.'

The findings showed that sentences containing NS were processed much faster than those with OSP but only when the NS was pragmatically resolved in favour of the matrix subject. According to Carminati, this evidences that the PAH is motivated by universal pragmatic principles in terms of the accessibility or informativity of referring expressions (Ariel 1990, 1994), since such violations do not entail grammatical problems but rather pragmatic ones.

In unambiguous contexts, adult monolingual speakers could disregard the PAH preferences, while in ambiguous contexts they generally tended to respect them. Thus, the PAH strategy is said to be more flexible in cases of non-ambiguity where both NS and OSP are appropriate. Disambiguation can be in principle provided by features of gender, person and number.

In sum, the PAH does not predict grammaticality since it is not a core-grammatical rule, but rather a processing hypothesis. Since the processing preference is context-dependent, the PAH is motivated by general cognitive mechanisms predicting felicity of a linguistic expression in context. The extralinguistic use of context is the distinctive characteristic of the so-called pragmatic meaning (Ariel 1990). Carminati (2002: 252) affirms that although the PAH operates on the basis of syntactically-encoded prominence relations in the discourse, 'the constraints it imposes are violable, in the sense that it dictates preferences, not (un)grammaticality' (see also Tsimpli & Sorace 2006). This accords with Ariel (2017: 476), who states that 'pragmatic violations create dispreferred (but grammatical) utterances'.

2.4.4 Salience

In order to account for the pragmatic factors that determine subject antecedents, prominence, referent salience or accessibility of the entity in the discourse context come into effect. According to Ariel's (1990, 1994, 2001, 2012) accessibility theory, referring (anaphoric) expressions convey degrees of accessibility of mental representations: the less accessible a referent is, the more elaborate the referential marker used by the speaker. In other words, the more information (i.e. semantic content and phonological material) an anaphoric expression contains, the less ambiguous it is and preferably retrieves a less salient referent (i.e. a less prominent one and/or with more competitors). Reduced referring expressions display a bias towards highly salient/prominent antecedents, i.e. highly explicit referents, which are more activated or accessible in people's minds. On the other hand, fuller referring expressions prefer to establish coreference to less salient/prominent antecedents (see also Hendriks 2003). This view is represented in the referential form hierarchy suggested by researchers such as Givón (1983), Ariel (1990) and Gundel et al. (1993) and it can be delineated in the following simplified cline:

Null > Pronoun > Demonstrative > Full Noun Phrase

The forms further to the right are used for less salient referents and vice versa (see Kaiser & Trueswell 2008). Thus, as Almor (2000) also remarks, there exists an inverse relation between referent salience and explicitness of anaphoric form. In language production and comprehension, selecting the appropriate referring expression and successfully identifying the antecedent are tasks whose ease of processing is determined by the best combination of computational cost and discourse function of the expression. The process reflects a balance between amount of information activated by the referential form in working memory and its discourse function. Memory cues are provided by anaphors[9] activating semantic information in working memory to identify referents (Almor 2000).

---

[9] The term 'anaphor' denotes any expression referring to an already established discourse entity (an antecedent).

The accessibility of a given antecedent in memory is achieved by any linguistic or extralinguistic factor that attracts attention on a particular entity increasing its prominence and thus the activation of its mental representation (Almor 2000). Subject position, specifically, confers salience on the entities of reference, 'such that the speaker can assume that the listener is more likely to focus attention on things in subject than nonsubject positions' (Arnold, Bennetto & Diehl 2009: 140). As Miltsakaki (2001) also argues, the grammatical role of subjecthood is the strongest salience factor in Greek. Apart from a prominent syntactic position, recency of mention is another factor which makes information salient since it is related to working memory resources (Gibson 1998; Arnold et al. 2013). Distance from previous mention of the antecedent, in particular, is correlated with potential ambiguity in referent identification (Givón 1983). Prada Pérez (2009) also affirms that probability to induce overt subject forms is increased as distance from previous mention becomes greater. Chances of potential ambiguity in interpretation are further increased in contexts where there is competition with other possible referents (Givón 1983, 2016; Ariel 1990; Arnold & Griffin 2007; Arnold et al. 2009; Lozano 2009; Blackwell & Lubbers Quesada 2012).

The inventories and functions of subject pronouns in NS and non-NS languages affect monolingual speakers' pronoun assignment to antecedents when handling pronominal ambiguity. Sorace and Filiaci's (2006) study on AR revealed that monolingual native Italian speakers generally assigned NS to subject antecedents in backward anaphora sentences, in line with PAH. However, in forward anaphora sentences, preferences of NS assignment were equally divided between matrix subject and matrix complement (see also Tsimpli et al. 2004). As the authors state, 'it seems that the pragmatic plausibility, topicality and accessibility (in terms of recency of presentation) of the complement all converge in overriding the PAS[10] and its bias against non-subject referents' (Sorace & Filiaci 2006: 358).

---

[10] It stands for Position of Antecedent Strategy, which is the same as PAH.

Salience is a complex phenomenon constrained by interaction of multiple syntactic, pragmatic and cognitive factors. Anaphoric forms can be sensitive to different factors to varying degrees (e.g. Miltsakaki 2007, 2011; Kaiser & Trueswell 2008). A reduced referring expression, e.g. a NS, can be used even when the referent is not salient insofar as it can be identified. The referential properties of NS, thus, do not force coreference to (preverbal) subjects or to the most salient antecedent (Carminati 2002; Blackwell & Lubbers Quesada 2012; Clements & Domínguez 2016; Frana 2017). According to Lubbers Quesada and Blackwell (2009: 120), 'switch focus of attention can occur with null subjects as long as other factors (e.g. discourse context, the interlocutor's preexisting knowledge, verb morphology) disambiguate the reference'.

NS used in sequential sentences to refer to different referents (i.e. in TS) are disambiguated through inferences and morphology. However, if the discourse involves only third-person referents of the same number and gender, with more than one potential antecedent in the interlocutors' common ground, verb morphology does not help disambiguation. It follows that third-person reference expressed via minimal forms, which are morphologically underspecified (i.e. NS), is inherently ambiguous since resolution depends on contextual and other morphological cues.

NS are normally used when the intended referent is the most salient entity at a given point in the discourse; therefore, antecedent saliency guides anaphoric reference and interpretation. Relevant factors in consonance with this pragmatic rule -except for morphology- are context, mutual knowledge of interlocutors and lexical semantics (see also Blackwell & Lubbers Quesada 2012).

2.5 Discussion

Referential subjects (LS, OSP, NS) as referring expressions contribute to information structure and discourse coherence while their particular form relates to the use and interpreting effects. In particular, NS and OSP in Greek and Spanish are neither in free variation nor in a strict complementary distribution. Since there is certain degree of overlap but also a qualitative distinction between the two pronominal forms, their occurrence is variable and thus not entirely predictable.

Pronominal subjects encompass a combination of two different aspects. One is related to language-specific core-grammatical properties, which licence NS and OSP. Subject pronouns are constrained by purely formal properties, i.e. syntactic (uninterpretable) features. The other aspect is associated with their discourse-pragmatic properties, which define their function in language as discourse-trackers (Pinto 2014). The use of subject pronouns is constrained by pragmatic conditions and discursive factors (i.e. interpretable features), related to the sentential information structure or pragmatic structure of the clause (e.g. Tsimpli et al. 2004; Lozano 2006b).

NS and OSP distribution, therefore, crucially depends on contextual conditions: some contexts promote the use of NS and others the use of OSP. Lack of compliance with discourse-pragmatic rules for subject use, despite not triggering ungrammaticality, may lead to lack of discourse coherence, which in turn may cause processing difficulties for an addressee. Discourse-level anaphoric resolution is thus challenging since it requires evaluation of different grammatically correct structures within the respective discourse context, a task which is taxing for working memory. Thus, pronominal AR is one of the most complex phenomena of natural language (Prentza & Tsimpli 2012; Shin & Cairns 2012).

Due to its complexity, the pragmatic distribution of subject pronouns is said to be acquired late in L1 acquisition, never mastered in a monolingual native-like fashion in L2 acquisition and to be susceptible in L1 attrition. These observations have led to a wealth of research as well as the formulation of the Interface Hypothesis, which is set forth in the next chapter.

3. Research framework

3.1 The Interface Hypothesis

Due to the fact that in NS languages the distribution of subject pronouns depends on the context, their distribution is considered to be located at the crossroads between syntax and pragmatics or discourse. These crossroads, or areas where syntax meets other linguistic modules, are the so-called interfaces. They are said to involve phenomena which are acquired late in L1 acquisition as well as to be vulnerable areas in bilingual acquisition, L1 attrition and L2 learning. This is because phenomena at interfaces require integration of different levels of linguistic knowledge.

For Sorace (2011, 2012) the term 'interface' is a descriptive device referring to syntactic structures that are sensitive to conditions which have to be satisfied so that the structure can be grammatical and/or felicitous. A grammar is compartmentalised into discrete modules: internal modules, such as lexicon, phonology, syntax, morphology and semantics; and external modules, such as discourse or pragmatics, referring to the meanings that a sentence has in a specific context. The different modules are connected by interfaces, i.e. areas of cross-modular integration of linguistic information. Internal interfaces consist of an interaction between only linguistic modules. In external interfaces, both linguistic and non-linguistic information is utilised. Interaction of language with non-linguistic cognitive systems through interfaces is illustrated by arrows in Figure 3.1.

Sounds (articulatory perceptual system)	PF	Lexicon		LF	Meanings (conceptual intentional system)
		Phonology	Syntax		
		Morphology	Semantics		

Context (discourse/pragmatics)
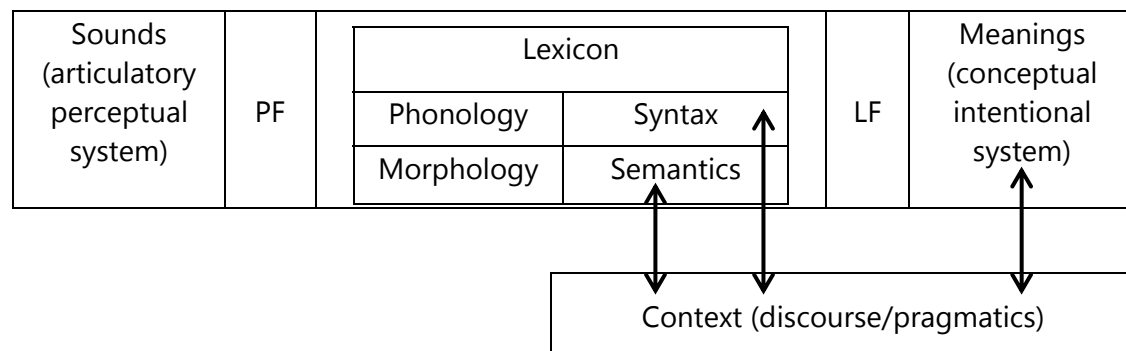
Figure 3.1. Internal and external interfaces (adapted from White 2009)

Interfaces are expected loci for difficulties, delays, grammatical erosion, arrested development or fossilization given their increased complexity compared to properties within narrow syntax (Sorace 2011; Tsimpli 2014). Narrow syntax, i.e. syntactic properties in a narrow sense or core-syntactic knowledge, is knowledge concerning structures that require only syntactic computations, e.g. the availability of NS. Crucially, 'in null subject languages, null subjects are syntactically licensed but their distribution is pragmatically determined' (Sorace & Filiaci 2006: 341). Narrow-syntax structures are acquired early in monolingual and bilingual acquisition, fully learned in L2 learning and remain stable in L1 attrition, unlike interface structures, which are more problematic (Sorace 2011; Tsimpli 2014). In words of Sorace & Filiaci (2006: 340) formulating what is known as the Interface Hypothesis (IH),

> [...] narrow syntactic properties are completely acquirable in a second language, even though they may exhibit significant developmental delays, whereas interface properties involving syntax and another cognitive domain may not be fully acquirable. This hypothesis has been found to have wider applicability in other domains of language development, such as bilingual L1 acquisition, L1 attrition, language breakdown and diachronic change.

Montrul (2011: 592), among others, contends that 'ultimately, every single utterance we utter involves discourse and must be read off at all linguistic interfaces', suggesting that no structures require solely syntactic computations and that demarcating separate interfaces cannot always entail clear-cut distinctions. Along similar lines, Domínguez (2013) claims that all grammatical structures necessarily involve interfaces (internal and/or external); consequently, satisfaction of interface conditions is always required as means to achieve convergence (i.e. full interpretation in minimalist terms). The question of whether there are structures involving only syntactic computations has caused much debate (see Montrul 2011; Sorace 2011; Domínguez 2013). In this regard, Sorace (2012: 213) suggests removing a rigid distinction between core syntax and interfaces and instead allowing 'for a range of interface conditions, graded according to their computational complexity and their dependence on extra-linguistic factors'.

Internal interfaces are considered to be less unstable than external ones. The external interface of syntax-discourse/pragmatics concerns pragmatic conditions of contextual appropriateness. This means that it requires additional processing resources while violations of its conditions result in a gradient of acceptability (optionality) and not ungrammaticality (Tsimpli & Sorace 2006; Sorace & Serratrice 2009; Tsimpli 2014).

Mastery of pronominal subjects demands not only the correct setting of the NSP in traditional principles-and-parameters terms, but also coordinating the pragmatic interface conditions that regulate the felicitous use of subjects in discourse (Sorace & Filiaci 2006; Tsimpli 2014). Due to the pragmatic nature of the feature of [±TS], OSP are characterised as being located at the syntax-discourse/pragmatics interface. This interface-based knowledge, depending on the discourse context, is predicted to be vulnerable and reveals persistent areas of optionality in language contact situations. Optionality is understood as 'the availability of two different forms to represent the same structure and express the same meaning' (Domínguez 2013: 78) as reflected in the '(in)consistent' linguistic behaviour of bilinguals or L2 learners when contrasted with the monolingual native speakers' performance (Tsimpli & Dimitrakopoulou 2007: 216). Optionality may also exist in monolingual native speakers (Sorace 2012). As regards interface phenomena, Sorace (2011: 12) notes that the speaker needs to have the following:

a. knowledge of the structure and of the mapping conditions that operate within interface components, and
b. the processing principles that apply in the real-time integration of information from different domains

As Tsimpli (2014) also explains, interface phenomena are by definition simultaneously sensitive to requirements of language-internal and language-external domains while real-time processing constraints on the integration of their properties are demanding both in quantitative and qualitative terms. Syntax-discourse/pragmatics interface phenomena regarding subject pronoun expression and interpretation have been consistently shown to be problematic in bilingual situations and processing (e.g. Montrul 2004a; Serratrice, Sorace & Paoli 2004; Tsimpli et al. 2004; Montrul &

Rodríguez-Louro 2006; Sorace & Filiaci 2006; Rothman 2009a; Sorace, Serratrice, Filiaci & Baldo 2009; Iverson 2012; Domínguez & Arche 2014; Kaltsa et al. 2015; Montrul & Sánchez-Walker 2015; Clements & Domínguez 2016; Montrul 2016a; Georgopoulos 2017). Acquisition of grammatical properties related to knowledge of discourse/pragmatic factors requires exposure to input that unambiguously provides the relevant constraints (Rothman 2009a; Tsimpli 2014). Insufficient knowledge of these interface properties leading to vulnerability in grammar is defined as 'residual optionality' in L2 acquisition, 'protracted indeterminacy' in bilingual L1 acquisition and 'emerging optionality' in L1 attrition (Sorace 2011: 5).

As far as person is concerned, Lozano (2009) showed that potential deficits with discursive features like [±TS], located at the interface, are selective since they do not affect the whole pronominal paradigm. The third-person singular referring to animate entities seems to be more vulnerable compared to the other persons, which appear relatively robust in the L2 context. In other words, the deictic use of pronouns, i.e. first and second person, and inanimacy (e.g. third-person neuter in Spanish) seem to be less vulnerable at the interface than the anaphoric use of the pronouns, i.e. third person animate (Lozano 2009).

The difficulty that interfaces present, as reported in bilingual speakers' performance, can be explained following two different lines of reasoning. The first one accounts for optionality assuming underlying deficiencies in the mental representation of the language. The second line considers the processing resources, necessary for simultaneous integration and application of information from different domains of linguistic knowledge, and their potential cost due to bilingualism. Thus, as regards the sources of instability at interfaces, optionality can be viewed through (a) the representational account and (b) the processing resources account.

3.2 The Representational Account

The representational account views optionality as weakening or underspecification of formal features instantiated in the mental grammar because one of the grammatical systems affects the other (Tsimpli et al. 2004; Tsimpli 2007; Tsimpli 2011). This account relies on the distinction between interpretable and uninterpretable features. 'An interpretable feature that is specified in L1 in a particular syntactic structure will become unspecified due to the absence of a similar interpretable feature in L2 in the same syntactic context' (Tsimpli et al. 2004: 263). As a result, there is a process of grammar restructuring in the bilingual speakers' mental grammars in order for them to adapt their mental representations of a certain construction in a way which fits both L1 and L2 linguistic systems (Pinto 2014).

On these grounds, it has been found that bilinguals of one NS language (e.g. Italian) and a non-NS language (e.g. English) tend to overgeneralise the use of OSP to contexts of TC. This is explained on the basis of weakening of pragmatic constraints or discourse-interpretable features regulating NS and OSP because of the knowledge of the non-NS language, which possesses the most economical syntax-discourse/pragmatics interface system for subject pronouns.

In the context of L1 attrition, Tsimpli et al. (2004) conducted a study on the referential pronominal system in L1 Greek and Italian in near-native speakers of L2 English. Distribution and interpretation of OSP was predicted to show L1 attrition effects because it is regulated by the interpretable features of [TS] and [Focus], which become underspecified due to language contact. In particular, underspecification of these interpretable features occurs because of the absence of similar interpretable features in English in the same syntactic contexts. Specifically, OSP specified for [+TS] and [Focus] in NS languages become underspecified due to attrition. Therefore, OSP interpretation does not necessarily indicate a shift in subject referent or focus in speakers' L1 and thus manifests more 'ambiguity' (Tsimpli et al. 2004: 263).

The Italian near-native speakers of English completed an off-line interpretation task involving backward and forward AR. In forward anaphora, the near-native group performed similarly to monolingual-native Italians linking the NS to the matrix

subject. In resolving OSP, while the monolinguals preferred a third referent, the near-natives did not manifest a clear preference, but rather 'multiple ambiguity', which confirms the predicted attrition effects regarding the overt form (Tsimpli et al. 2004: 273). In backward anaphora, the experimental group mostly linked the NS to the matrix subject, whereas the controls opted for either the matrix subject or the complement. The authors hypothesise that the embedded clause may be treated as non-finite by the near-native group, which would make the matrix subject necessarily control the embedded NS. As for the OSP in backward anaphora, both groups preferred to link it to the matrix complement; however, there was a significant difference between the two groups' percentages of preference for the matrix subject as a possible referent. This result confirmed the prediction that OSP interpretation would display attrition effects since the near-native speakers allowed an OSP to be interpreted as a continued topic significantly more than the control group.

The crosslinguistic effects behind overuse of OSP in TC in bilinguals are seen as obeying economy factors and having a specific directionality. The (uni)directionality can be predicted by the structural composition of the linguistic systems depending on their complexity. It is expected to be from the pragmatically less complex language (the non-NS language) to the other one (the NS language), regardless of whether it is L1 or L2 (Sorace 2004, 2011). This assumption accounts for the overuse/overacceptance of OSP in bilingual speakers in contexts where NS are favoured by monolinguals (Serratrice et al. 2004; Sorace et al. 2009). The overuse or redundancy of OSP, although not syntactically incorrect, is considered pragmatically odd, inappropriate or infelicitous.

In Tsimpli et al. (2004), there was no evidence of attrition in either Greek or Italian speakers of English in syntactic aspects related to subjects that are due to the parameterisation of purely formal (i.e. uninterpretable) features. Attrition effects were found in aspects of subjects' distribution and interpretation which are relevant to syntactic and pragmatic (i.e. interpretable) features. The study demonstrates the selective pattern of areas that can be affected in language contact situations and, particularly, the permeability of the interface of syntax-discourse/pragmatics, which gives rise to inconsistent grammars or optionality (see also Tsimpli 2007).

Differences from monolinguals at the level of knowledge representation can be presumably detected only in bilinguals speaking a language pairing in which one language has a complex setting and the other does not. This is seen as a limitation of the representational account since it applies only to language combinations with different pragmatic settings. Sorace and Serratrice (2009: 199) argue that in language pairs where both languages instantiate a complex setting, 'no underspecification, and therefore no optionality, would be expected for these combinations'. This suggests that differences between monolingual and bilingual performance are not predicted in subject distribution in a bilingual situation involving two NS languages.

Interestingly, overextension of the scope of OSP has been shown to obtain in bilingual situations not only when a NS language comes in contact with a non-NS language, but also with another NS language; hence irrespective of L1/L2 pairing and typological similarity (Bini 1993; Margaza & Bel 2006; Lozano 2006a, 2017; Sorace et al. 2009; Tammer 2016; Georgopoulos 2017). Sorace et al. (2009) found over-acceptance of redundant OSP in Spanish-Italian bilingual children living in Spain. This may indicate that 'the scope of the overt pronoun in Spanish is actually wider than in Italian' (Sorace et al. 2009: 474). Apart from potential crosslinguistic differences among NS languages, the aforementioned findings suggest that processing factors play a non-trivial role in real-time production and comprehension, hence the processing resources account (Sorace 2011), which aims at explaining these results.

3.3 The Processing Resources Account

The processing resources account perceives optionality as caused by processing-related difficulties that are intrinsic to bilingualism at the level of language use (Sorace & Filiaci 2006; Sorace et al. 2009; Sorace & Serratrice 2009; Sorace 2011). The latest version of the IH argues that sources of optionality are difficulties with 'the processing principles that apply in the real-time integration of information from different domains' (Sorace 2011: 12). Cognitive capacity, such as working memory, is more likely to be less efficient in performing tasks, since language operations such as retrieval and computation may present degrees of deficiency. Sorace and Serratrice (2009) assert that difficulties at external interfaces can be explained considering

general processing and attentional deficiencies that arise when a bilingual speaker, dealing with more than one internal grammar, tries to coordinate syntax with contextual discourse-pragmatic information. Montrul (2011) also argues that the complexity of tasks used to examine speakers' interface knowledge should be considered before making claims on the interface phenomena.

In the context of L2 acquisition, Sorace and Filiaci (2006) used a picture verification task to examine forward and backward AR in ambiguous intrasentential contexts in near-native L2 Italian speakers with L1 English and monolingual Italian adults. The results showed that the near-native speakers manifested a native-like strict PAH in their NS resolution preferences in both forward and backward anaphora. However, they showed indeterminacy in the OSP resolution preferences, which significantly differed from those of the native control group. Specifically, the near-native performance revealed a more flexible PAH for the resolution of OSP, which allowed coreference with both a subject and a non-subject antecedent, irrespective of contextual ambiguity.

These findings suggested that the OSP may correspond to a default or unmarked form in L2 Italian. Additionally, since the language pair in the study included a non-NS language (L1 English) and a NS language (L2 Italian), the role of transfer or L1 interference could not be disregarded. Thus, the L1 was viewed as an extra factor contributing to the tendency of the OSP to be used as the default form, even in near-native speakers. According to the authors, 'the combination of sub-optimal processing resources and crosslinguistic influence may increase the magnitude of the effects' (Sorace & Filiaci 2006: 346). Therefore, crosslinguistic influence is regarded as a reinforcing factor and not the only cause of this phenomenon (see also Sorace 2005). It was concluded that the near-native participants had acquired the NS grammar and knowledge of the PAH; however, they possibly did not have 'the necessary processing resources to integrate multiple sources of information' (Sorace & Filiaci 2006: 361). Sorace (2014: 72) explains inefficient processing as default reliance on OSP and comments on the processing account as follows:

> The processing explanation [...] is not an alternative to the linguistic explanation; rather, it is an account based on consideration of the interaction between linguistic and general cognitive factors. These factors are not only 'memory limitations or heavy cognitive load' [...], but are rather important aspects of executive function that affects the way in which speakers incrementally build central coherence in anaphora resolution.

The previously-mentioned considerations are based on findings which clearly indicate that there is more than crosslinguistic influence in the L2 performance of some speakers. In particular, overuse of OSP has been found in combinations of two NS languages in different bilingual domains, as will be shown in the following sections.

3.4 Overuse of overt subjects in pairs of null subject languages

3.4.1 L2 learning

3.4.1.1 Spanish L1 - Italian L2

Bini (1993) examined the basic grammatical properties of the NSP in Spanish L2 learners of Italian including two proficiency levels: beginner and intermediate. She elicited spontaneous speech data from oral interviews, which she considers adequate in methodological terms because, while speaking, participants do not have the time to reflect much upon their production, so the rules of their interlanguage are more easily disclosed. The participants employed redundant OSP in their L2 Italian in contexts where a NS would be favoured in both Spanish and Italian. All three grammatical persons were considered in the study. According to the author, while the pattern arisen did not coincide with either the L1 or the L2 use of OSP, it started to approach more the L2 distribution (i.e. Italian) after six weeks of further L2 instruction. Overproduction of OSP was an unexpected finding given that the language combination comprised two typologically similar NS languages.

Scrutinising her data, Bini (1993: 136) mentions that overuse of the first person *io* 'I' may be (partly) due to an 'egocentric tendency' of the speakers. In other participants, OSP could substitute for verbal morphology, which may be incorrect or absent as it is

often the case in lower-proficiency learners' performance. She also observes that the learners (over)use the pronoun in order to compensate for hesitation, to 'fill' a pause and thus to provide time for planning what to say next without interrupting the discourse. In other cases, speakers wanted to ensure that the subject of the clause was identified if it had been previously introduced through verbal inflexion only. Bini emphasises the following fact: Since learners are conscious of their limited L2 proficiency, they prefer using overt forms, i.e. not omitting the pronoun, so that they do not have to rely on correct use of verbal morphology, which they are still learning. OSP overuse does not disappear in later stages of L2 learning; however, it is observably mitigated, which suggests that this phenomenon is not permanent.

3.4.1.2 Greek L1 - Spanish L2

Margaza and Bel (2006) investigated L2 acquisition of Spanish OSP in intermediate and advanced L1 Greek learners living in Greece. The authors' hypothesis was that the learners would have positive transfer from their L1 either only in the syntactic domain, exhibiting misuse of subjects in the pragmatic domain; or both in the syntactic domain and at the syntax-pragmatics interface, manifesting a native-like behaviour. The Spanish learners and Spanish controls completed a cloze test, in which they had to select missing subjects in different discursive contexts, and a free production task consisting of writing a short narrative.

The results of the tasks showed that both L2 groups produced NS in both matrix and embedded clauses, which evidenced knowledge of the parametric option for subjects in Spanish. However, intermediate learners overused OSP in contexts requiring NS when compared to advanced and monolingual native speakers. Although there was no evidence of verbal inflection errors, it was argued that intermediate learners 'express the subject to reinforce verbal morphology' (Margaza & Bel 2006: 96).

Margaza and Bel (2006) ponder that if there was transfer of L1 pragmatic knowledge to L2 Spanish, overuse of OSP would not occur, since speakers would prefer the NS option as expected in their L1 Greek. Thus, there was apparently positive L1 transfer only in the syntactic domain and not in the pragmatic one. Since this was the case of

the intermediate level only, it was inferred that proficiency level plays an important role affecting the use of NS and OSP. The study indicates, as in Bini (1993), that misuse (overuse) of OSP, despite the pair of NS languages, is a transitory phenomenon in the L2 development of a NS language.

Lozano (2006a, 2018) studied three levels of adult Greek learners of L2 Spanish living in Greece (intermediate, lower advanced, upper advanced) using a written offline contextualised acceptability judgement task on AR. The contexts included TC (a single antecedent requiring a NS), contrastive-focus and emphasis. Even at end-states, the L2 learners showed tolerance of pragmatically redundant OSP, significantly more than Spanish controls, but less so than the less proficient L2 groups. These results also confirmed that L2 learners' sensitivity to the AR pragmatic constraints improves along with proficiency.

Georgopoulos (2017) investigated a written corpus of Greek and English adult intermediate and advanced learners of Spanish focusing on pronominal resolution of third-person anaphoric subjects in Spanish. All subjects were coded for felicity/infelicity depending on information status. Unpragmatic subjects included overexplicit (redundant) or underexplicit (ambiguous) subjects. Overexplicit were those subjects that appeared in TC, i.e. pronouns or noun phrases referring to the subject of the previous clause. A noun phrase was also considered overexplicit in contexts of two different-gender referents, in which an OSP would be the felicitous option. Underexplicit subjects were those involving NS in TS. Additionally, an OSP in any context with several same-gender referents was considered underexplicit due to same-gender ambiguity.

The Greek intermediate group produced significantly more overexplicit subject forms than Spanish controls, but this was not the case in the advanced Greek group. It was concluded that Greek (and English) learners' performance becomes more native-like as proficiency increases. The native-like performance of the Greek advanced learners is evidence against the IH's prediction of optionality at this interface structure even at near-native proficiency levels. The findings were in line with other AR studies attesting native-like performance in advanced/near-native learners (e.g. Judy 2015).

Few cases of underexplicit NS were also reported, while most cases of illicit subjects involved overexplicit (overt) subjects. It was noted that native Spanish speakers were also occasionally redundant in subject use. The results showed that there is L1 influence in the production of anaphoric subjects, since the Greek-speaking groups performed significantly better than the English-speaking ones. Thus, differences between the two groups of learners could be straightforwardly explained by crosslinguistic influence. The performance of the Greek intermediate group producing overexplicit subjects as well as the use of underexplicit NS, although minimal, were not further explored.

3.4.2 Adult Bilingualism

3.4.2.1 Moroccan Arabic L1 - Spanish L1

Bel and García-Alcaraz (2015) used an AR acceptability judgment task to examine two NS languages, Moroccan Arabic and Spanish. Monolingual Arabic oral performance in a sentence interpretation task on antecedent preferences was similar to that reported for Spanish (Alonso Ovalle et al. 2002; Filiaci 2011) corroborating the PAH only regarding NS resolution, since OSP were more inconsistent. Early sequential Arabic-Spanish bilinguals (adolescents) and Arabic adult L2 learners of Spanish participated in the study as well as a Spanish control group. The experiment was administered as a paper-and-pencil task in which the continuation of completely ambiguous sentences was judged according to a four-value Likert scale.

Overuse of OSP in TC was found in bilinguals in matrix-embedded order but no over-acceptance of NS was detected in any condition, while no deviance with either NS or OSP was revealed in the L2 group. The differences attested between bilinguals and L2 learners were attributed to different rates of input exposure. The authors conclude that 'the preference of null pronouns in same-reference intra-sentential contexts seems to be a robust and steady phenomenon among different null-subject languages, even typologically distinct (Romance and Arabic), whereas overt pronouns are more unpredictable' (Bel & García-Alcaraz 2015: 225).

3.4.3 Child bilingualism

3.4.3.1 Spanish L1 - Italian L1

In order to shed light on the crosslinguistic effects and their directionality in child bilingualism, Sorace, Serratrice, Filiaci and Baldo (2009) conducted a large-scale study considering the dominant societal language. The participants were English-Italian bilingual children living in the UK and in Italy and Spanish-Italian bilingual children living in Spain (aged 6-7 and 8-10) as well as monolingual children and adults. An acceptability judgment task was used to test knowledge of NS and OSP in [+TS] and [-TS] contexts in English and in Italian based on a short animated story involving cartoon characters. Examples from the Italian task are (18) and (19) for [-TS] and [+TS] respectively. The participants were asked to select the 'better' option. In (18) the felicitous answer is Donald's response, while in (19) the felicitous answer is Mickey's response (Sorace et al. 2009: 467).

(18)   (Minnie and Daisy in the foreground; Mickey and Donald in the background)

       Minnie: pro Sono caduta!

            '[I]'ve fallen!'

       Donald: Minnie ha detto che pro è caduta.

            'Minnie has said that [she] has fallen.'

       Mickey: Minnie ha detto che lei è caduta.

            'Minnie has said that she has fallen.'

(19)   (Minnie and Daisy in the foreground; Mickey and Donald in the background)

       Minnie: Daisy è caduta!

            'Daisy has fallen!'

       Donald: Minnie ha detto che pro è caduta.

             'Minnie has said that [she] has fallen.'

       Mickey: Minnie ha detto che lei è caduta.

            'Minnie has said that she has fallen.'

The English results revealed that the monolingual children and adults and the older bilingual children all scored at ceiling accepting the sentences with OSP in both in [+TS] and [-TS] and rejecting those with NS. The younger bilingual children scored only marginally less accurately than the other groups and this difference was significant merely in the acceptance of NS in [+TS]. Thus, the syntactic constraint for English was easily learned and influence from Italian was minimal.

The Italian results were different. The adult Italian speakers did not always reject the pragmatically inappropriate option, i.e. they accepted a limited number of redundant OSP in [-TS] (see also Carminati 2002; Tsimpli et al. 2004; Sorace & Filiaci 2006 for similar results). This provides confirmatory evidence that monolingual native speakers' judgments in AR are not categorical since they reveal some tolerance towards redundant forms especially in contexts where there is no ambiguity risk.

That said, adult Italian speakers' scores were significantly 'better' than those of all children groups. The younger children selected OSP in [-TS] contexts significantly more often than the older children, who also preferred OSP significantly more often than the adults in this condition. The younger bilingual children living in the UK selected OSP in [-TS] significantly more often than those living in Italy or Spain, revealing greater crosslinguistic influence from English.

Interestingly, the bilinguals of Spanish-Italian, i.e. two NS languages, unexpectedly performed much like the English-Italian bilinguals choosing OSP in [-TS] significantly more often than the monolingual Italian children. The over-acceptance of OSP, indicating lack of sensitivity to redundancy, implies that the status of such form may be a default to which speakers resort when efficient coordination of different factors involved in AR falls short. This may be due to the developmental stage in which children are and to the fact that processing operations may be somewhat hindered because of the effort of having two language systems in bilingual children. In Spanish-Italian bilinguals, over-acceptance of OSP increases with age, which may suggest a high sensitivity to Spanish pronoun interpretation. Thus, a difference in the scope of OSP in Spanish compared to Italian was also considered as a plausible scenario (Sorace et al. 2009), but was not explored any further.

In [+TS] contexts, monolingual adults and children selected OSP more often than both English-Italian and Spanish-Italian bilinguals, with the bilingual children of both language pairs accepting more NS in the [+TS] condition than the monolingual Italian children and adults. Thus, the bilingual children's problems were not only with redundancy but also with ambiguity in accepting more NS in contexts that favour OSP in Italian. Sorace et al.'s (2009) findings lend support to the claim that ambiguity is sooner overcome than redundancy by monolingual children. Ambiguous NS in [+TS] were explained as the result of 'the bilingual children's more taxed processing resources' (Sorace et al. 2009: 475).

The findings indicate that crosslinguistic influence is not the only factor at work in the interpretation of pronouns. Of particular interest is the over-acceptance of both OSP and NS by bilingual children, regardless of language combination. It is stated that inappropriate OSP in [-TS] 'give rise to redundancy without compromising the assignment of pronouns to their antecedents' whereas inappropriate NS in [+TS] contexts 'give rise to ambiguity and prevent the successful identification of the pronoun antecedent' (Sorace et al. 2009: 473).

In sum, divergent behaviour at external interfaces was attested in bilingual speakers due to influence from the L1/L2 grammar and/or real-time processing cost. The extra burden of processing due to bilingualism was considered responsible for (at least some) deviations from end-state monolingual norms in areas requiring integration and coordination of syntactic and contextual information. Unidirectionality of crosslinguistic influence from the non-NS to the NS language was confirmed but still over-acceptance of ambiguous NS in the Italian performance by all bilingual groups could not be sufficiently explained in unidirectionality terms. Although this phenomenon is ultimately overcome, it questions the directionality of crosslinguistic effects. The findings in Sorace et al. (2009) provide grounds to explore crosslinguistic differences between NS languages in more depth.

3.4.4 Discussion

The pattern emerging from these (and other) works shows parallel effects of optionality as found in different situations of language contact, i.e. attrition (e.g. Tsimpli et al. 2004), bilingualism (e.g. Sorace et al. 2009), L2 learning (e.g. Sorace & Filiaci 2006; Georgopoulos 2017). The languages in contact are characterised by overextended use or acceptance of OSP by L2 learners and bilinguals compared to monolingual native speakers. The phenomenon manifests itself as being impermanent in the case of L2 learning.

With regard to attrition, Chamorro, Sorace and Sturt's (2015) work on L1 Peninsular Spanish AR with English as the other language showed that L1 re-exposure reduces attrition effects, thus frequency and recency of re-exposure are important factors to be considered. As attrition effects were not shown to be irreversible, L1 attrition seems to be related to processing of interface structures rather than to a permanent change in speakers' L1 knowledge representations.

Overuse of OSP attested in combinations of two NS languages, even though an apparently surmountable phenomenon in L2 learning, is in fact a counter-intuitive finding, given the 'facilitative' language pairing (Judy 2015). The substantial overlap in two NS languages in pronoun realisation patterns would be expected to result in positive influence with no overgeneralisation of OSP use in bilinguals.

Since overextension of the scope of OSP cannot be attributed to parametric crosslinguistic influence and especially with regard to Sorace et al.'s (2009) findings, it may be explained as a compensatory strategy occurring on account of bilingualism itself (Sorace et al. 2009). The ability to efficiently coordinate and process syntactic and pragmatic/contextual information in real-time performance is affected by the simultaneous activation of the two linguistic systems, which moves attentional resources away from other (linguistic) tasks (Sorace 2011, 2012). Due to attentional resources allocation, bilinguals are said to experience an increased processing burden, which results in default strategies such as the (over)use of OSP (Sorace 2011).

In this regard, Tsimpli (2011) observes the distinction between 'learner default' and 'linguistic default'. The former is the OSP employed by bilinguals to compensate for inefficiency in syntax-pragmatics mapping, while the latter is the NS as the default or unmarked form in NS languages. The possibility of crosslinguistic differences in NS languages regarding distribution of NS and OSP is a crucial factor which must be also considered in language-contact situations.

3.5 Crosslinguistic differences in null subject languages

Although NS languages can be analogous regarding the general universal principles that govern NS vs OSP distribution, they are not identical with respect to the particular scope of these pronouns (Carminati 2002; Gürel 2006; Sifaki & Sitaridou 2007; Sorace et al. 2009; Prada Pérez 2010, 2018; Filiaci 2010, 2011; Sorace 2011, 2012; Iverson 2012; Mayol 2012; Filiaci, Sorace & Carreiras 2013; Pinto 2014; Bel & García-Alcaraz 2015; Judy 2015; Duarte & Soares da Silva 2016; Tammer 2016, a.o.).

Judy (2015) conducted a study on near-native Spanish speakers of L1 Farsi, a NS language with pronoun distribution similar to Spanish. In felicity judgment tasks, tolerance of NS in TS contexts was higher in Farsi than in Spanish, indicating a wider scope of the NS in Farsi compared to Spanish. Farsi and Spanish are typologically distinct NS languages, which might explain this difference (cf. Bel & García-Alcaraz 2015). However, even in typologically similar NS languages, it is argued that typological relatedness seems to be overestimated (Pinto 2014). Some typical cases of closely-related Romance language pairings with well-attested differences in subject pronoun distribution are those between European and Brazilian Portuguese (e.g. Mayol 2012; Tammer 2016) as well as between Peninsular Spanish and Catalan (Prada Pérez 2009, 2010, 2018). Even between different varieties of Spanish, there are important dissimilarities in pronominal subject distribution, with typical case the difference between Caribbean and non-Caribbean varieties (Lipski 1994, 2012; Montrul 2004b; Ordóñez & Olarrea 2006; Mayol 2012; Domínguez 2013).

Italian and Spanish, closely related Romance languages, exhibit differences in the scope of OSP, as suggested in Sorace et al. (2009); hence, they cannot be considered equivalent regarding the properties of their pronominal system. Antecedent preferences of personal pronouns varied across these two languages as also demonstrated in Filiaci (2011). In particular, she showed that the Spanish pronouns *él* and *ella* lack a restriction on the possibility to be associated with prominent discourse antecedents. This was a sign of relative weakness of Spanish personal pronouns compared to their strong Italian cognates (*egli* and *ella*), indicating that the two NS languages do not share the same discourse-pragmatic properties. Given the differences found between NS languages on their discourse-pragmatic properties, crosslinguistic influence cannot be ruled out as a possible source of non-target performance (Filiaci et al. 2013), thus the representational approach cannot be a priori rejected (Pinto 2014).

3.6 The Vulnerability Hypothesis

Contrasting the views of the IH, the Vulnerability Hypothesis (VH) was put forward by Prada Pérez (2018) to formally analyse cross-linguistic influence. According to this proposal, linguistic complexity resides in variability and not in interface conditions. Variability is viewed as the availability of two or more forms for a specific grammatical pattern, whose alternation is not categorical. There is a variability continuum or permeability hierarchy predicting vulnerability in distributions of linguistic phenomena surfacing in crosslinguistic effects in language contact situations depending on complexity. Specifically, 'the more variable a distribution is the more susceptible to cross-linguistic influence it will be' (ibid: 4), entailing that bilinguals resemble monolinguals in the most categorical variables and differ in the least categorical variables. The extent to which a speaker exhibits cross-linguistic influence also depends on individual and social factors related to language contact intensity.

The author tested these predictions examining monolingual and Spanish-Catalan bilingual speakers' performance in Spanish and bilingual's performance in Catalan focusing on speech connectivity, which refers to co-referentiality and TAM in subject expression. Speech connectivity was found to be the most categorical variable in

Spanish, but distribution of Spanish pronominal subjects can be also interpreted as lying at the syntax-pragmatics interface. Thus, according to the IH, speech connectivity is vulnerable due to interface conditions, whereas VH would predict that the aforementioned variable is rather invulnerable due to its categorical nature. The study showed that the categorical distribution of speech connectivity was not affected by language contact as attested in the bilingual performance, in line with the VH and against the IH predictions. Prada Pérez (2018) argues against the IH emphasising the problem of interface identification and the potential location of structures at more than one interface, a phenomenon which is difficult to measure. By contrast, the VH proposal is based on the concept of variation, which is feasibly measurable. Thus, variable vs categorical discourse behaviour in monolingualism may predict vulnerability to crosslinguistic influence in the bilingual domain.

3.7 Overuse of null subjects

Perhaps the main shortcoming of the IH comes from the observation that construction and interpretation of all linguistic expressions demand contextual or pragmatic information interacting with syntax (Rothman 2009a; Montrul 2011; Domínguez 2013). In this respect, Sorace (2011: 25) argues that 'instead of a rigid dichotomy, it may be more appropriate to differentiate structures on a gradient according to the type of conditions they need to satisfy and whether they are closer to the "strongly biasing" syntactic end or to the "weakly biasing" contextual end'.

Clements and Domínguez (2016) bring to the fore the overextension of the scope of NS, on which discussion has been inconclusive. Their assumption is that both NS and OSP manifest degrees of complexity at the interface because both forms can be used in [±TS] contexts in a NS language such as Spanish. They claim that NS have been inaccurately characterised as unproblematic in L2 acquisition because of their apparently simple information structure, contrary to their overt counterparts. OSP are specified for [+TS] and [Focus] and therefore carry pragmatic meaning which places them at the syntax-discourse/pragmatics interface, while NS lack these features. According to the IH, due to their pragmatic nature, OSP are persistently difficult, while this is not the case for NS. However, NS are also used in TS when referent

identification is possible though means of salience; therefore, interface conditions also apply. Research has revealed overuse and/or over-acceptance of NS in bilingual situations, suggesting that NS are also complex in referential terms; consequently, their pragmatic properties may also present instability.

3.7.1 L2 learning

English L1 - Spanish L2

On the basis of the complexity of NS, Clements and Domínguez (2016) employed two interpretation tasks in order to examine knowledge of NS and OSP in L2 Spanish in different pragmatic contexts involving [±TS]. The participants were advanced speakers of L2 Spanish with L1 English as well as monolingual controls of Spanish speakers from Spain and Mexico. The results corroborated the authors' predictions on potential difficulties encountered by L2 speakers regarding NS distribution, since the participants demonstrated non-native-like preferences in over-accepting NS as compared to monolingual controls. Directionality of crosslinguistic effects as explained in Sorace (2004, 2011), i.e. the language with the more restrictive interface system (English) influencing the language with the less economical system (Spanish), was challenged since it could not account for the results.

Clements and Domínguez (2016) explain such findings by arguing as follows: Although NS acquisition is not problematic in L2 Spanish, indeterminacy as to whether a NS or OSP is more appropriate leads learners to 'avoid using a pronoun altogether which appears as though a null subject has been intentionally selected' (Clements & Domínguez 2016: 24). Additionally, in some cases classroom instruction may encourage the use of NS as a default form for the sake of preventing overuse of OSP if learners' L1 is English. The authors hold that NS can be easily used as default because the high accessibility of their referent as well as other means, such as verbal morphology, help recover their content.

Rothman (2009a) also found overuse of both OSP and NS in the performance of L2 learners of Spanish with L1 English. He hypothesised that intermediate and advanced learners have knowledge of the NSP but encounter problems with the distribution of

NS and OSP constrained by discourse/pragmatic factors. He sought to verify whether these problems are unidirectional (overuse of OSP) or bidirectional (overuse of both OSP and NS). The experiments included interpretation and translation/production tasks. The L2 learners had acquired the syntax of NS early and straightforwardly compared to OSP pragmatic distribution associated with the interface. Although there were difficulties in the acquisition of OSP, these were ultimately surmounted. Part of the findings was in line with the IH, namely the intermediate L2 group displayed lack of sensitivity to overuse of OSP. Interestingly, the same group also under-used OSP in production, that is, they overused pragmatically odd NS in environments with ambiguous verbal morphology (i.e. third person singular and plural) obscuring subject identification. This challenges the claims by Sorace and Filiaci (2006) that L1 influence and L1/L2 differences in processing strategies favour the (over)use of OSP as a default because of the syntax-discourse/pragmatics interface conditions posing difficulties even in near-natives.

Rothman (2009a) endorse the syntax-before-discourse position corroborating the additional complexity involved in integrating syntactic and pragmatic information related to pronominal subject distribution and to L1 influence from English. However, the unidirectionality of effects was not supported because of the attested overuse of both OSP and NS. Rothman (2009a: 967) explains this finding as a 'bidirectional target-deviancy', which makes subject pronouns be in a state of free variation, unrestricted by the relevant external interface conditions in developing L2 Spanish. These pragmatic delays, however, are temporary and eventually overcome at more advanced levels of proficiency. Thus, the interface between syntax and discourse/pragmatics, although problematic in L2 acquisition, does not necessarily imply an area of fossilization in L2 Spanish. The complexity of the distribution of NS, in addition to that of OSP, as attested in works such as Clements and Domínguez (2016) and Rothman (2009a), has been also revealed in preceding studies (e.g. Pérez-Leroux & Glass 1997, 1999; Rothman 2007a; Rothman & Iverson 2007).

Montrul and Rodríguez Louro (2006) examined the acquisition of the syntax-discourse/pragmatics interface in L2 Spanish questioning whether discourse-pragmatic constraints are acquired along with morphosyntactic properties of the NSP

or at a later stage in L2 development. L2 speakers (intermediate, advanced and near-native) with L1 English participated in an oral production task narrating the story of *Little Red Riding Hood* based on a picture-story booklet.

The results showed that all learners had acquired the relevant morphosyntactic properties, but the intermediate group produced significantly more redundant overt subjects in TC and more agreement errors than the advanced, the near-native and the control groups. The advanced group also overused overt subjects but to a lesser extent compared to the intermediate group. According to the authors, this deviance emerged because the intermediate learners' Spanish was still constrained by the discourse-pragmatic properties of their L1. As for the NS, their use was considered illicit when the story line and the agreement morphology on the verb indicated a switch of subject reference but the speaker did not produce an overt subject. The near-native and especially the advanced learner group overused illicit NS (5% and 8.4% respectively), while the intermediate and the native-speaker groups did not.

The fact that the advanced learners overproduced both overt subjects and NS is evidence both for and against 'unidirectionality', which predicts overuse only for overt subjects arising from underspecification of the feature of [+TS]. NS are considered to lack this feature, thus no NS overproduction is predicted, contrary to Montrul and Rodríguez Louro's (2006) findings. As regards NS overuse, the authors state that since 'this type of error decreases by the near-native level of proficiency, it might not end up as a prototypical feature of indeterminate and incomplete grammars to the extent that the opposite error pattern -overproduction of overt subjects- appears to be' (Montrul & Rodríguez Louro 2006: 414). In sum, L2 learners of Spanish demonstrated late acquisition of the distribution of Spanish referential pronouns due to the additional discourse-pragmatic complexity, which seems to be gradually overcome. It remained unestablished whether overuse of NS could be 'experimental noise', a performance error, or a true sign of linguistic deficit (Montrul & Rodríguez Louro 2006: 415).

3.7.2 Heritage acquisition

English L1 - Spanish L1

Montrul (2016a) examined the rate of NS and overt subjects and their discourse/pragmatic distribution in same reference and switch reference contexts using the same oral production elicitation task as in Montrul and Rodríguez Louro (2006). The participants were school-age bilingual children, young adult HS, adult Mexican immigrants and two monolingual age-matched controls from Mexico. In addition to the overuse of overt subjects, which was expected, both child and adult bilingual speakers produced pragmatically illicit NS at the same rate as the Mexican monolingual school-age children. This finding showed that the bilingual children presented a similar developmental pattern to that of their monolingual peers, which however was enhanced by the bilingual context in which they lived. Interestingly, the same tendency was also detected in adult HS, indicating a developmental link between the two groups. The author claims that 'while bilingual school-age children show developmental and bilingualism effects, the adult HS' grammars are the result of bilingualism, reduced input in childhood, and possibly input effects from relatives who may undergo attrition' (Montrul 2016a: 244).

3.7.3 Adult bilingualism / bidialectalism

Brazilian Portuguese L1 - European Portuguese L2

Tammer (2016) investigated the distribution of NS and OSP in two very closely-related, mutually intelligible, Romance variants: European (EP) and Brazilian Portuguese (BP). In EP, embedded OSP establish coreference in line with PAH. However, in BP the person-verb agreement paradigm has been losing its morphological distinctions causing an increase in frequency of OSP and hence yielding free variation of OSP and NS. As a result, embedded pronoun subjects manifest a mixed pattern, thereby presenting optionality. Tammer compared groups of adult HS of BP who acquired L2 EP as children in Portugal, BP natives who acquired L2 EP in Portugal as adults and monolingual controls of both Portuguese variants.

In the AR experiment, which was based on Sorace & Filiaci's (2006) task, EP and BP controls significantly differed in OSP resolution. Both groups linked the overt form to the object, but BP speakers allowed for coreference with the subject significantly more than the EP speakers. In the narratives task, there was a significant difference in the rates of OSP, with BP producing them significantly more than EP speakers.

As regards the bilingual/bidialectal groups, the IH was taken into account predicting that both groups would overuse and over-accept OSP in EP due to crosslinguistic effects from BP. Indeed, the L2 learners accepted OSP in contexts where EP speakers preferred NS. In production, they also used significantly more OSP in EP than the controls. The HS patterned with EP monolinguals in both EP production and interpretation of NS, without manifesting the default strategy of overusing OSP, contra the IH. Crucially, in BP production both target groups overused NS manifesting crosslinguistic effects from EP to BP, which was not anticipated by the IH, since NS are not considered an interface structure. Potential crosslinguistic effects due to bilingualism were expected to surface as overuse of OSP and not the reverse, which was in fact attested in the study. In sum, the target groups performed similarly in production in both EP and BP, while significant differences were found in interpretation in both modes, which were explained on the basis of age of acquisition.

3.8 Ambiguity in overuse of null subjects

Montrul (2004a) investigated crosslinguistic influence from English to Spanish examining pronominal preferences in two groups of adult HS of Spanish (intermediate and advanced) and a group of monolinguals. Oral production was elicited by requesting participants to narrate in Spanish the story of *Little Red Riding Hood*. As far as discourse-pragmatic appropriateness is concerned, NS were coded as licit in contexts of same referent (TC) and as illicit in contexts of change of referent (TS) or contrastive focus. In the latter cases, NS are assumed to make the context ambiguous or unclear. By contrast, overt subjects were coded as redundant in TC and as correct in TS or emphatic contexts.

The results showed that the intermediate heritage group produced more overt subjects than NS, while the advanced group was more accurate than the intermediate group in both NS and overt subjects' expression. Crucially, however, both groups were found to use more pragmatically inappropriate NS than overt subjects, signifying a challenge for the unidirectionality predictions. Some examples provided from the intermediate group may show cases where, although a NS is used in TS, the referent is not actually ambiguous, as in (20) (Montrul 2004a: 133).

(20) *Caperucita Roja salió a ir a la casa de su abuelita con una canasta de comida porque ∅ estaba, ∅ iba a visistarla porque *∅ estaba enferma.*

'Little Red Riding Hood went out to go to her grandmother's house with a basket of food because ∅ was, ∅ was going to visit her because *∅ was sick.'

In (20), the knowledge shared by the interlocutors because of the familiarity with the story and the given context are sufficient to assign the correct referent to the NS. Further, the antecedent of the inappropriate NS, i.e. the grandmother, is brought into focus as a prepositional object and as pronominal direct object. Thus, identification of the intended referent without particular difficulties questions the purported pragmatic inappropriateness of the NS (see comments in Liceras et al. 2010; Lubbers Quesada 2015). It seems to be the case that qualifying NS as ambiguous is not a straightforward matter in evaluating production data. As claimed by Sorace (2004), the overproduction of NS in this study must be considered with precaution as evidence against unidirectionality because of the insufficient 'information gap' between the participants and the investigator.

In response to this last point, Montrul and Rodríguez Louro (2006), who used the same methodology with similar findings, maintain that the familiarity with the story was an advantage of the methodological approach. This is because a well-known tale, such as *Little Red Riding Hood*, makes easier to control the vocabulary, precisely when groups of different levels of proficiency are involved. Moreover, this type of task permits observation of different grammatical properties. This is not possible in tasks with isolated sentences, which are not particularly useful for examining the relationship between grammatical elements and their discourse function.

The authors also invoke Lafond, Hayes and Bhatt (2001), who likewise report overproduction of NS in TS contexts in L2 Spanish by beginner, intermediate, advanced and near-native speakers. The task in Lafond et al. (2001) for data elicitation was totally different consisting of selecting the more appropriate of two provided responses according to participants' judgement in order to complete an unfinished dialogue. The learners demonstrated use of NS since early stages, but without discriminating discourse contexts, hence causing NS overgeneration. This behaviour appears to show a progressive decrease by the near-native level, which means that it is transitory.

An example from Lafond et al. (2001: 126) is (21), which illustrates the type of discourse contexts given to the L2 participants. They were asked to choose one of the two given answers to the final question of speaker A in the dialogue. The target option is (2), i.e. the overt subject *Beth* is expected to be mentioned because of TS.

(21)	*A: Hola John*	'Hi, John'
	B: Hola Ana	'Hi, Ann'
	A: ¿Te gustaría almorzar conmigo?	'Would you like to eat with me?'
	B: Sí, me gustaría. ¿Puede venir Beth también?	'Yes, I would. Can Beth come too?'
	A: Seguro, ¿a qué hora quieres ir?	'Sure, when do you want to go?'
	B: 1) Bueno, *∅ está en clase ahora. ¿Está bien a las 12:30?	'Is in class now. Is 12:30 ok?'
	2) Bueno, Beth está en clase ahora. ¿Está bien a las 12:30?	'Beth is in class now. Is 12:30 ok?'

It can be argued, however, that *Beth* is a previously mentioned referent, which is the only antecedent in third person singular in the dialogue (all other subjects-referents are in first and second person). The NS of the third person verb *está* 'is' cannot be assigned to any other referent. In other words, verb inflection may solely refer to Beth (person disambiguation), due to absence of competing referents. The above-mentioned NS, although not the most felicitous option, should not be considered illicit since it is not ambiguous within its context. More generally, a NS is not illicit if person, number, or gender may unambiguously identify its antecedent (see Carminati 2002).

Montrul and Sánchez-Walker (2015) investigated the development of subject expression in Spanish-English heritage bilingual children and adolescents in USA. They looked into NS and overt subjects' rates and distribution in narrative production (story-telling) of bilinguals of 6-17 years old as well as age-matched monolinguals from Mexico. The monolingual children were different from the adults in Montrul (2004a) in using significantly more NS in TS contexts. Monolingual children of all ages used infelicitous subjects (LS, OSP, NS) while 9 of them (of a total of 20) used ambiguous NS in TS, such as in (22) (Montrul & Sánchez-Walker 2015: 242).

(22) *Había una vez una niña que se llamaba Caperucita Roja. Y era como su cumpleaños y estaba bailando la niña porque *∅ le había regalado una capa roja y *∅ estaba bailando.*

'Once upon a time there was a little girl named Little Red Riding Hood. And it was like her birthday and was dancing the girl because *∅ had given her a red hood and *∅ was dancing.'

The bilingual children produced pragmatically illicit NS at the same rate as monolinguals. This finding, according to the authors, suggests that the participants had not fully developed pronoun reference in Spanish. Additionally, the bilingual children of all ages produced more redundant overt subjects than their monolingual peers. Subject expression was found to be a vulnerable area in child bilingual grammars. However, the patterns were the same as those of monolingual development with a different magnitude of the effects, i.e. higher in the bilingual domain. Some of the factors that contribute to the attested indeterminacy of pragmatic distribution of NS and overt subjects include influence from the dominant language (English), structural complexity of overt subjects in Spanish, integration of syntactic and pragmatic features, contact with other dialects, as well as age of onset of bilingualism (Montrul & Sánchez-Walker 2015: 244).

For Liceras et al. (2010), lack of ambiguity is the crucial condition for distinguishing pragmatically appropriate NS from 'illicit' ones. Thus, a more fine-grained analysis of potential ambiguities is possible by replacing the ambiguous NS with an OSP. For

instance, use of the OSP *ellos* 'they' in place of *∅ in (23) from Montrul (2004a: 133) would not make the sentence less ambiguous.

(23)   *[...] Mientras estaba dormido el lobo *∅ le llenaron el estómago con piedras, y la abuelita estaba lista pa coserle el estómago.*

'[...] While was sleeping the wolf, *∅ filled the stomach with stones, and the grandmother was ready to sew him the stomach.'

Liceras et al. (2010) showed that the ambiguity illustrated in (20) is also present in monolingual native Spanish. This was observed particularly in cases of ambiguous verb forms between first, second and third persons (see §2.3.7.2). It was argued that this type of NS distribution in L2 grammars cannot be fully explained by interface vulnerability accounts because of its occurrence in monolingual grammars too. The authors propose a distinction between two types of ambiguity: (a) NS yielding ambiguity which cannot be resolved by the use of an OSP; therefore, a noun phrase (LS) is required; and (b) NS in contrastive focus contexts or NS which violate the PAH.

With respect to ambiguity, Pinto (2014) reports overextension of the scope of NS in Italian in oral production (narratives) and interpretation (AR) in adult bilingual performance with L1 Italian and L2 Dutch (a non-NS language). She showed that NS were regularly used in TS by both monolingual and bilingual speakers in story-telling and in a picture verification task. A comparison between bilinguals and monolinguals revealed no differences between the groups. Since even monolinguals did not perform according to the 'norm', the author questions the validity of the assumed baseline for the distribution and interpretation of subject pronouns in Italian. In the qualitative analysis of the subjects as produced in oral narratives, she found that NS appeared in TS without causing ambiguity. Pinto (2014) observes that TS combined with omitted subjects appears in environments where the context offers cues to identify and interpret the NS. The speakers exploit other resources provided by the language to signal TS, such as (a) contextual cues, i.e. semantic information that identifies the referent and its discourse function, and (b) grammatical cues, such as agreement morphology on the verb or elsewhere, which renders OSP use pleonastic (see also Pinto 2013).

Taken together, interpretation and production data in Pinto (2014) shed light on the PAH concerning both monolingual and bilingual speakers in both AR and (semi)spontaneous oral production. As regards interpretation, since the test sentences were ambiguous and not contextualised, the participants relied on the standard processing strategy proposed by the PAH, which determines the possible antecedent based on its position in the construction. Thus, the PAH was found to be a default processing strategy applying in the interpretation of structures where the antecedent identification of the pronoun creates ambiguity. As for the production data, 'in the case of spontaneous speech the PAH plays a minor role; null subjects are frequently used and ambiguity is avoided by exploiting cues from the context' (Pinto 2014: 88). Thus, when there is no risk for ambiguity because of the given context, speakers perform with flexibility without resorting to the default interpretation but rather using information provided by contextual cues.

In earlier work, Paredes Silva (1993) also asserted that pronoun usage in subject position is constrained predominantly by discourse and that 'the more predictable the reference of the subject, the more likely it is to be omitted' (Paredes Silva 1993: 45). Morphologically ambiguous structures may or may not be contextually ambiguous; therefore, potential ambiguity should be evaluated at the morphological and the discourse level paying attention to context. Human inferential abilities and world knowledge are viewed as strong factors which play a part in determining interpretation of ambiguity (Pinto 2013, 2014; Wasow 2015; Winkler 2015).

3.9 The role of age differences in adulthood

Studies of language production and interpretation evidence that adult aging involves cognitive decline, such as limitations of short-term working memory, perception and attention span, which influence referential skills (see Hendriks et al. 2014 and references therein). Brain aging, according to Shafto and Tyler (2014), does not affect the most core aspects of language processes, but it entails cognitive aging and it is related to a consequent decline in aspects of language production. Much of speech comprehension, on the other hand, is generally well-preserved across the lifespan, despite the complex computations required to automatically access and construct

linguistic representations. In words of Shafto and Tyler (2014: 583), 'although aging is associated with specific impairments in language production, most comprehension abilities remain stable as we age'. Abrams and Farrell (2010) explain that age-related deficits on spoken language tasks concern particularly word-retrieval problems causing disfluency. This is because speech production requires activating and retrieving phonological information, a process which becomes slower and more effortful with aging. It has been also showed that age-related inhibition problems affect language production, giving rise to decreased coherence in spoken language related to producing more off-topic speech (off-target verbosity) (James, Burke, Austin & Hulme 1998; Arbuckle, Nohara-LeClair & Pushkar 2000). Although language comprehension is said to be relatively unaffected in elderly speakers, varying degrees of memory loss and hearing loss, which are also associated with aging, may undoubtedly affect processing abilities in interpreting oral language (Pelc 2001).

Restricting attention to referential skills relevant to the present study, expressing reference in discourse and resolving anaphoric pronouns are associated with domain-general cognitive systems, which determine the efficiency to process referential choices. Both comprehension and use of referential devices are cognitively demanding tasks and have been reported to decline with aging within adulthood; they are thus subject to age-related cognitive constraints (see Abrams & Farrell 2010; Hendriks et al. 2014; Ghaleh 2015; Kaltsa et al. 2015).

Hendriks et al. (2014) tested both production and comprehension of referential choices in Dutch children, young adults and elderly adults. The method used included a storytelling task and an anaphora comprehension task based on pre-recorded stories. Both tasks involved two same-gender characters. Their findings showed that the elderly speakers produced significantly more ambiguous pronouns in TS contexts than the younger adult speakers. On the other hand, in listening comprehension both young and old participants performed similarly in identifying the referent shift. It was thus observed that referential choices in comprehension did not exhibit significant differences between the two adult age groups. Hendriks et al. (2014: 404) conclude that 'elderly speakers have difficulty in keeping track of the structure of the discourse and determining the prominence of the referents in the discourse'.

Ghaleh (2015) also studied native speakers of New Zealand English in a series of AR experiments. She found that older adults were less accurate in interpreting anaphoric pronouns than younger adults due to differences in working memory capacity, which affects comprehension accuracy. The author also considered the possibility that comprehension may be mediated by a decline in inhibitory abilities responsible for the process of suppression of relevant information previously provided. It was overall suggested that 'both differences in working memory capacity and age-related decline in suppression abilities can affect discourse comprehension' (Ghaleh 2015: 154).

Age-related changes in language processing involving resolution of ambiguous anaphora were also reported in Kaltsa et al. 's (2015) study on monolingual Greek and bilingual Greek-Swedish groups (see details in §4.1.2). The older group in the AR experiments showed weak preferences in resolving both overt and null pronouns compared to the younger group. As to the role of chronological age, the authors confirm that 'age within adulthood plays a role in the tendency to discriminate between preferences in pronominal resolution overall' (Kaltsa et al. 2015: 281).

Schmitz and Scherger (2017) analysed production data from natural conversations testing adult HS of Italian living in Germany. The Italian performance of HS, contrary to monolinguals, showed that use of overt first-person pronoun increases at higher ages. This, however, was not the case for the third person. Such findings also point to age as a factor to be considered in assessing the production/omission of subjects.

3.10 Discussion

The findings confirming overuse of OSP in NS languages firmly suggest the need for a more in-depth approach when examining the formal and discourse-pragmatic properties of pronominal subjects in different languages. Even closely related NS languages may be different in the scope of their pronouns; hence potential crosslinguistic differences and their source should be identified and taken into account. In addition to the overuse of OSP even in combinations of two NS languages, the inverse pattern has been also found, that is, over-extension of the scope of NS, which still remains relatively unaddressed. For instance, in Sorace et al.

(2009) over-acceptance of NS involved all bilingual children groups, regardless of age, language pairing and community language. NS in TS are generally considered to be under-explicit, which leads to the question of whether and to what extent such instances of NS should count as illicit. Ambiguity is a key condition in this respect. Crucially, NS are felicitously used in TS contexts by both monolingual and bilingual speakers of NS languages (Liceras et al. 2010; Blackwell & Lubbers Quesada 2012; Pinto 2013, 2014).

As for the qualitative difference between OSP and NS overuse, Rothman (2009a: 967) states that 'overusing overt pronouns is not wrong per se. It is simply pragmatically odd. Worse, however, is the failure to use overt subjects when the discourse information does not provide an immediately identifiable/accessible subject'. Judy (2015: 171) also opine that 'using overt subjects in topic maintenance contexts is perhaps less infelicitous than using null subjects when overt subjects are required for contrastive focus of topic shift (i.e. a question of redundancy versus lack of crucial information)'. Sorace et al. (2009: 473) made similar claims on NS' overuse, which can be problematic since it may undermine communication. Inappropriate NS may result in ambiguity and hence processing complexity with consequences in communication. Since in production NS always occur in context, the point which needs to be foregrounded is how crucial the missing information of a NS is, whether it generates ambiguity and whether ambiguity is incidental or temporary or it remains unresolved. A more in-depth look into production data may offer insights into the (default) interpretation strategies in resolving contextless ambiguous anaphoric reference.

As an interface phenomenon, subject pronouns distribution and interpretation in NS languages is demonstrated to be an unstable phenomenon in language contact situations, with redundancy of overt subjects being a recurring characteristic of this instability. As far as ambiguity is concerned, due to the multiplicity of factors affecting NS use (e.g. presence or absence of morphological cues and/or competing referents), evaluating NS as (non-)ambiguous requires attention to particular contexts of use. Finally, the role of age at the time of testing is an important variable potentially affecting referential skills and ability of both producing and interpreting anaphoric reference in discourse. It is thus crucial that this variable is taken into account.

4. Greek and Spanish studies

4.1 Research studies in Greek

The present section provides a review of the most relevant studies conducted on Greek speakers with regard to the distribution and interpretation of subjects in Greek.

4.1.1 Monolingual Greek

One of the first studies on pronoun resolution strategies in Greek was conducted by Dimitriadis (1996) using the Centering Theory (Grosz et al. 1995). Available antecedents are viewed as 'forward looking centres'. The 'centres' of each utterance are potential antecedents for the next utterance obeying a discourse prominence ranking. Corpus data indicated that the distinct semantic role of NS and OSP results in a division of labour. NS tend to link to the available highest-ranked centres, i.e. salient/prominent antecedents, which are compatible in terms of grammaticality and agentivity. The referring preference of NS is not as strong as that of OSP, since OSP are rarely construed with the most prominent antecedent of the preceding utterance. 'The overt pronoun signals this merely by its presence, not by any featural or referential information that it provides' (Dimitriadis 1996: 2). The strong bias towards a less salient/prominent antecedent is thus considered to be part of the OSP' conventionalised meaning. However, the division of labour only partially depends on the lexical properties of pronominal forms since pragmatic factors allow flexibility in establishing anaphoric relationships; thus, AR is non-categorical.

Miltsakaki (2007) examined referential properties of NS and OSP in Greek in intersentential AR using written completion data elicited from a free sentence continuation task. She followed the claims of the semantic focusing hypothesis of Stevenson, Crawley and Kleinman (1994): the verb, due to its semantic interpretation, has focusing properties depending on the thematic role of the entities. Action verbs focus on the entity with the patient role independently of its syntactic realisation, i.e.

even if it is realised as the object of a verb. The monolingual Greek native speakers had to complete sentences such as (24) from Miltsakaki (2007: 3).

(24)   *O travmatioforeas kuvalise ton astheni. Ekinos...*
        'The stretcher-bearer carried the patient. HE...'

The results showed a strong tendency of the Greek speakers to continue the given discourse with reference to the entity with the patient role in object position. In a second similar study, the continuation after the first clause was completely free, that is, the beginning was not provided, so the participants could freely use a referent and a referring expression of their choice. When the continuation was with a matrix clause, reference to the object-patient was chiefly made by means of a strong pronoun or noun phrase. When the continuation was with an embedded clause, the object-patient was more often referenced with a NS. On the other hand, reference to agent-subject antecedents was accomplished using NS in both matrix and embedded clause continuations. Altogether, the most common continuation (44%) was with embedded clauses and null reference to the object-patient of the previous sentence, such as in (25) from Miltsakaki (2007: 4).

(25)   *I mitera filise tin kori, epiði ∅ perase tis eksetasis.*
        'The mother kissed the daughter, because [she] passed the exams.'

The findings show that a mapping between choice of linguistic expression and entity salience is a complicated notion. Miltsakaki (2007: 5) argues that 'entity accessibility within the boundaries of a syntactic sentence is primarily determined by the semantic relations that are established by the predicates of the main and embedded clauses'. It may also be the case that the Greek demonstrative is employed for object reference in matrix clause continuations in a comparable way to the use of a stressed pronoun in English, i.e. to mark TS, or to counteract parallelism effect. In matrix clause continuations, NS preferentially linked to the subject referent while strong pronouns referred to the object referent. Finally, word order (i.e. the position of subject) did not significantly play a role in the participants' choice of reference in continuations.

Argyri and Sorace's (2007) study investigated the potential crosslinguistic influence from English to Greek in the use and interpretation of subjects in Greek based on the claim that these are syntax-pragmatics interface structures considered to be vulnerable. The participants were simultaneous bilingual eight-year-old children living in Greece and in the UK as well as monolingual controls. The interface constructions involved distribution of NS vs OSP in [-TS] contexts and distribution of preverbal vs postverbal subjects in wide focus contexts. The narrow-syntax constructions involved the use of preverbal vs postverbal subjects in *what*-embedded interrogatives and the placement of objects. The methods consisted of an elicited production picture task and an off-line acceptability judgement task, whose design was based on questions and answers. Examples of the items used for elicited production of NS or OSP in [-TS] and the forced-choice acceptability judgment task are (26) and (27) respectively.

(26) Question: *Jati pije i Eleni_j sto periptero?*
      'Why did Eleni_j go to the kiosk?'
      Expected answer: *Epiði ∅_j ithele na aγorasi efimeriða.*
      'Because [she]_j wanted to buy a newspaper.'

(27) Question: *Jati pije o Jorγos_j sto vivliopolio to proi?*
      'Why did George go to the bookshop this morning?'
      Puppet A: *Epiði ∅_j ithele na aγorasi ena vivlio.*
      Puppet B: *Epiði aftos_j ithele na aγorasi ena vivlio.*
      'Because he_j wanted to buy a book.'

Monolingual Greek children aged 7;5-9;7 years and adults always produced a NS in the [-TS] condition. However, in the acceptability judgment task the Greek children did not consistently accept the NS utterance in this condition, unlike adults. This is an interesting finding pointing out difference in performance as manifested in production and interpretation modalities. It also shows that monolingual children, at least up to the age of 9 years old, may accept redundant OSP for topic reference, i.e. for reference to salient subject antecedents.

Prentza and Tsimpli (2012) explored AR of pronominal ambiguity in Greek monolingual adults applying two conditions: pragmatically neutral and pragmatically biased conditions. The test sentences presented contexts of forward anaphora with the linearly first clause being either matrix or embedded. The latter clauses were introduced with the subordinators 'when/while' and 'because' as well as the coordinators 'and' and 'but'. The results showed that, although pragmatically plausibility played a role, NS were preferentially attached to subject referents by monolingual Greeks and OSP to object referents. Crucially, the aforementioned correlation was stronger in the case of OSP referring to object antecedents than in NS referring to subject antecedents, even in cases of pragmatic biases. The authors conclude that the form which is more underspecified for reference, i.e. NS, is consistent with the default TC interpretation and that the OSP is strongly associated with object antecedents in Greek.

Papadopoulou et al. (2015) conducted a study on resolution of intrasentential ambiguous NS and OSP anaphora in Greek children and adults. The children participants were classified into groups aged 6-6;3, 7-7;8, and 10-11;2. Two experiments examined coordination of grammatical and discourse-pragmatic information, required for NS and OSP ambiguity resolution in sentence interpretation. One experiment focused on the resolution of NS and the other experiment on the resolution of OSP. The data elicitation method consisted of a self-paced listening picture verification task and the test sentences were segmented as shown in (28).

(28)   *O papus / miluse / ðinata / ston egono tu / otan / ∅ ðjavaze / ena vivlio.*
        'The old-man spoke loudly to his grand-child when (he) read a book.'

The adults consistently preferred to associate NS with matrix subjects and OSP with matrix objects. In the OSP condition, the object was chosen significantly faster than the subject referent, while in the NS condition the subject was not chosen faster than the object referent. Children, on the other hand, showed a different pattern. OSP were resolved in an adult-like fashion by the oldest children, but this was not the case for the two younger groups who also linked OSP to subjects. NS manifested a 'U-shaped' development, since the two younger groups performance was similar to

adults, but the oldest child group selected the object antecedent more often than all other groups. Consequently, there was no clear NS resolution pattern in Greek at the age of 10-11 years. The authors claim that OSP carry the discourse feature of [+TS], whereas NS are underspecified regarding discourse features, thus do not directly select a preferred antecedent allowing more reference possibilities. It was also stressed that the feature of [+TS] is not a grammatical feature, hence (un)grammaticalities do not obtain in any resolution routines. AR is regulated by pragmatic principles involving degrees of preferences rather than categorical distinctions between grammaticality and ungrammaticality, as also asserted by other researchers (e.g. Carminati 2002; Tsimpli & Sorace 2006; Sorace & Filiaci 2006; Rothman 2009a; Shin & Cairns 2012; Kaltsa et al. 2015; Georgopoulos 2017).

4.1.1.1 Definiteness in anaphora resolution

Mastropavlou, Katsiperi, Fotiadou, Fleva, Peristeri and Tsimpli's (2014) study looked into AR in adult Greek considering also the factor of definiteness in relation to the syntactic positions of the antecedents. The factors examined included the form of anaphoric expression (null vs overt), the syntactic position of the antecedent (subject vs object) and their interaction, i.e. if the reduced (null) forms prefer a prominent antecedent (subject) and the strong (overt) forms select a less prominent antecedent (object). Additionally, it was assumed that definiteness could enhance or diminish antecedent prominence. The structures examined were (a) a matrix clause with a definite subject and a definite object, which is considered to be the baseline; (b) a matrix clause with a definite subject and an indefinite object, and (c) a matrix clause with an indefinite subject and a definite object. As regards (c), the authors explain that when the subject is weakened because of indefiniteness, the object is not expected to be definite, which renders this structure marked. The matrix clauses were followed by an adverbial clause, which presented a NS or an OSP, and the participants chose its antecedent in an online self-paced listening and picture-matching task. An example is shown in (29) (Mastropavlou et al. 2014: 11).

(29) *I/Mia γramateas voithuse ti/mia nosokoma otan afti/iði eγrafe ena γrama.*
'The/A secretary was helping the/a nurse when she/already was writing a letter.'
Pjos eγrafe to γrama?
'Who was writing the letter?'

In the OSP condition, there was a significant effect of definiteness in the structure with indefinite subject. The structure with indefinite subject and definite object marked the slowest reaction times, while the structure with both definite subject and object was found to be the fastest one. In the NS condition, definiteness also had a significant effect with the same pattern of reaction times. With regard to the antecedent preferences, the object antecedent was significantly preferred in the OSP conditions. However, NS were not significantly chosen as referring to subjects in the conditions of definite subject-definite object and definite subject-indefinite object. Moreover, embedded NS in structures with matrix indefinite subjects and definite objects were linked to objects significantly more than with subject antecedents.

In sum, it was demonstrated that definiteness manipulations affected antecedent preferences in the NS condition. Preference for the object antecedent in the NS condition was intensified when the subject of the matrix clause was indefinite, marking also an increased processing difficulty as indicated by listening and response times. Since interpretation of sentences with OSP was not affected by definiteness alternation, the study suggests that definiteness affects AR only in cases of NS sentences, where no preference was attested, i.e. cases of high ambiguity. In sentences of low ambiguity, i.e. those involving OSP, the markedness of the overt form seems to be stronger than the definiteness of potential antecedents.

4.1.2 Heritage and attrited Greek

Kaltsa, Tsimpli and Rothman (2015) examined NS and OSP resolution in two adult bilingual populations, Greek HS and Greek L1 immigrants with Swedish as the other language (a non-NS language). Two monolingual control groups also participated, namely a younger group and an older group, in order to match HS and immigrants' age respectively. Antecedent preferences for NS and OSP as well as reaction times

were examined in the study, focusing also on the role of age at the time of testing and the factor of bilingualism. The two experiments consisted of a self-paced listening and sentence-picture matching task. The target sentences presented a matrix clause with two referents and an adverbial clause with a NS or an OSP, based on the experiment of Papadopoulou et al. (2015). The participants indicated the picture which better described the sentence that they had listened.

With regard to OSP resolution, it was found that all groups, monolingual controls included, showed a preference for object antecedents. Differences were found only regarding the subject antecedent, which was chosen significantly more by the two bilingual groups as compared to the respective age-matched controls. As for the NS resolution, all groups preferentially anchored NS to subject antecedents. However, the older monolingual group did not show a clear preference for either the subject or the object antecedent. In addition, there was a significant difference between the HS and the younger monolingual group in that the former showed weaker preferences for either subject or object antecedents in comparison to their monolingual peers.

The age factor was separately scrutinised by dividing all monolingual and bilingual participants in only two groups: (a) a young group, consisting of the HS and the younger monolinguals; and (b) an old group, formed by the immigrants and the older monolinguals. In the OSP condition, both groups clearly preferred association with the object antecedent. The only significant difference was on the percentage of subject antecedent preferences, which was higher in the case of the old participants when compared to the young group. In the NS condition, both groups showed a preference for the subject antecedent, which was stronger in the young group and weaker in the old group. Crucially, however, the two groups were significantly different in their matching decisions as regards the object and the third-referent antecedents in this condition. The old group linked the NS to the object antecedent and the third-referent significantly more often than the young group. Thus, the old speakers displayed weaker preferences for both pronoun forms, i.e. they allowed association of subject and object antecedents with both NS and OSP. The young participants' preferences were stronger by comparison, as they tended to select the subject over the object antecedent in NS resolution. This finding reveals an age effect

in pronominal resolution, which suggests that age in adulthood is a determinant for antecedent preferences. Particularly, in NS resolution the subject antecedent preference seems to become weakened with age. The authors conclude that 'interface vulnerability is open to factors other than attrition or bilingualism as such' (Kaltsa et al. 2015: 281), with age manifesting itself as one of them.

The effect of bilingualism as a factor on its own was also examined by separating the total of participants into two different groups depending on their linguistic background, namely a bilingual and a monolingual group, with no age distinction. The OSP was shown to be more often correlated to the object antecedent by both groups. However, the bilingual group chose the subject antecedent significantly more times than the monolingual group. Turning to the NS, both groups favoured the subject antecedent and there was no significant difference between the groups. This evidences that 'bilingualism seems to trigger differences in pronominal resolution of overt pronouns only' (Kaltsa et al. 2015: 282).

In conclusion, the authors state that crosslinguistic influence could not be the only cause for the findings since similar preferences of the OSP referring to the object antecedent were also found in the older monolingual group. In contrast, preferences for the NS resolution were stable. Thus, a purely processing account cannot fully explain the phenomena, which could be examined considering the linguistic properties of the NS and OSP and particularly those of the OSP, which appears to be more vulnerable.

4.1.3 Greek: Summary

Taking into account the Greek studies, NS in general are preferentially linked to subject antecedents whereas OSP favour coreference with object antecedents (Dimitriadis 1996; Miltsakaki 2007; Prentza & Tsimpli 2012; Kaltsa et al. 2015; Papadopoulou et al. 2015) in line with PAH. This division of labour is not mutually exclusive (categorical) since both pronominal forms (NS, OSP) may access a prominent or less prominent intended antecedent depending on the discourse context. Moreover, the strength of coreference between NS and subject antecedents

seems to be weaker than the strength of coreference between OSP and object antecedents (Dimitriadis 1996; Prentza & Tsimpli 2012; Mastropavlou et al. 2014; Papadopoulou et al. 2015). It should be also noted that in Mastropavlou et al. (2014), while the OSP was consistently attached to object antecedents, NS were shown to be more unstable without favouring subject antecedents in the conditions of definite subject and in/definite object in the matrix clause. In Greek, OSP generally mark TS and therefore lead to strong preferences to be assigned to non-salient referents. NS, although tending to relate to subject antecedents, seem to be more flexible hence potentially more ambiguous than OSP (Dimitriadis 1996; Mastropavlou et al. 2014).

The Greek studies also point to crucial factors to be considered in investigating pronominal resolution patterns. Miltsakaki's (2007) study showed that sequences of two matrix clauses may reveal different preferences in establishing reference relations than sequences of matrix and embedded clauses. Mastropavlou et al. (2014) highlighted the role of definiteness in AR. Kaltsa et al. (2015) brought into focus the factor of age at testing, since the older monolingual speakers manifested different resolution routines from the younger speakers. There are also important differences in the resolution routines between forward and backward anaphora contexts (Tsimpli et al. 2004). The Greek data suggest differences in pronoun resolution preferences between Greek and Spanish, since in Spanish OSP do not seem to demonstrate a strong bias in favour of non-salient antecedents, in contrast to Greek, as will be seen in the next section.

4.2 Research studies in Spanish

4.2.1 Monolingual Spanish

Spanish has been extensively studied with respect to pronominal resolution strategies in monolingual and language contact domains. The present section reviews some of the most relevant studies in monolingual Spanish.

4.2.1.1 Peninsular Spanish

Alonso-Ovalle, Fernández-Solera, Frazier and Clifton (2002) examined preferences in pronoun assignment in intersentential contexts in Peninsular Spanish following the predictions of PAH. The data elicitation method was a written questionnaire. An example from Alonso-Ovalle et al. (2002: 166) is (30).

(30) *Sara abrazó a Teresa. ∅/Ella está emocionada.*
       'Sara embraced Teresa. ∅/*She* is excited.'

In more than 73% of the time the NS was linked to the subject antecedent showing a clear coreference of the null form to the preceding subject. By contrast, the OSP was attached to either the subject or the object antecedent around 50% of the time, revealing a random behaviour of the overt form in terms of antecedent preferences. Thus, pronominal resolution in Spanish only partially supported the PAH since NS showed a clear tendency towards TC selecting subject antecedents, but OSP remained neutral between TC and TS assigning antecedents at chance level contra the PAH. The findings suggested crosslinguistic differences between NS languages (i.e. Spanish and Italian) on subject pronoun distribution.

Filiaci (2011) ran a series of self-paced reading experiments in Peninsular Spanish and Italian in order to compare the two NS languages in the performance of adult speakers. The experiments were based on the PAH and Carminati's (2002) methodology. The results replicated Carminati's (2002) findings for Italian, corroborating the attested strict division of labour between NS and OSP. However, in Spanish this division of labour was not well defined. The association between OSP

and TS was found to be weak in Spanish compared to Italian. The results indicated presence of cross-linguistic differences between the two languages on the interpretation of OSP, which was asymmetrical in the two languages in that the OSP in Spanish was apparently weaker than in Italian. Since the author employed exactly the same methodology in both languages, her comparison endorsing crosslinguistic differences between Spanish and Italian is well founded. Similar findings regarding Peninsular Spanish are reported in Filiaci, Sorace & Carreiras (2013).

Additionally, Filiaci (2011) examined the role of the relative amount of homophony (hence ambiguity) among different forms in the verbal paradigm of Spanish and Italian with respect to use of NS and OSP. The weakness of the OSP bias attested in Spanish compared to Italian may be related to the higher relative frequency of verb forms that are not uniquely marked for person (i.e. homophonous) in Spanish, rendering ambiguity (see §2.3.7.2). Thus, OSP may be used to explicitly indicate the person features that are missing and cannot be recovered otherwise. If this is the case, then there is a loss of the structural bias of Spanish OSP (should they have one). Filiaci (2011) conducted an AR experiment in which, apart from variables of anaphora (NS vs OSP) and antecedent (subject vs object), she manipulated the verb tense. The verbs could be either in present indicative (with unambiguous verbal morphology) or in present subjunctive (with ambiguous verbal morphology in both languages). Additionally, she included a third tense condition, namely the imperfect, which is ambiguous in Spanish but not in Italian. The experiment was a phrase-by-phrase self-paced reading task. An example of these sentences is (31) from Filiaci (2011: 167).

(31) *Beatriz_i ha obtenido un ascenso de Carmen_j, aunque ella_i/∅_j sea inexperta para el nuevo trabajo.*
      'Beatrice_i has obtained a promotion from Carmen_j, although she_i/∅_j is inexperienced for the new job.'

The results did not support the hypothesis that OSP are used with ambiguous verbal morphology, thus there was no evidence that the potential structural bias of the OSP was lost or weakened. Likewise, there was no evidence of preference of OSP over NS regardless of antecedent position due to morphological ambiguity in verb inflection.

Moreover, according to Filiaci (2011), Spanish NS and OSP seem to be roughly interchangeable in terms of processing cost. Filiaci et al. (2013: 16) conclude that 'Spanish overt pronouns are relatively insensitive to syntactic prominence compared to Italian pronouns but also to null subjects in both languages'.

While Alonso-Ovalle et al. (2002) and Filiaci (2011) report similar findings on Peninsular Spanish, opposing results were reported by Chamorro et al. (2015). In their AR study using intrasentential matrix-embedded anaphora with temporal clauses, NS lacked a consistent pattern of antecedent preferences whereas OSP were strongly associated with the object antecedent.

4.2.1.2 Chilean Spanish

Chilean Spanish was the focus of Callahan, Nicol, Love and Swinney (2007) on processing and interpretation of NS and OSP in intersentential contexts investigating the effects of discourse prominence in adult speakers. The task employed was a self-paced reading task, which presented two-sentence passages with one or two mentioned referents and a NS, an OSP, or a name as the embedded subject. Examples of the test sentences of the study are shown in (32) and (33):

(32) *Juan trabaja cada día hasta las ocho. Juan cree que ∅/él/Sam debería ir de vacaciones.*
'Juan works every day until eight o'clock. Juan thinks that ∅/he/Sam should go on vacation.'

(33) *Sam trabaja cada día hasta las ocho. Juan cree que ∅/él/Sam debería ir de vacaciones.*
'Sam works every day until eight o'clock. Juan thinks that ∅/he/Sam should go on vacation.'

The participants identified the subject of the embedded clause in a multiple-choice questionnaire. The results indicated that both NS and OSP were interpreted as referring to the discourse prominent referent. Thus, there was no variation found according to pronoun type. However, reading times for the OSP clauses were longer

than for the NS clauses. When there was an interpretative bias, both NS and OSP were interpreted as referring to the more plausible referent, regardless of discourse prominence. This is an indication, according to the authors, that pronoun biases interact with plausibility biases in determining final conscious interpretation. The reading time responses in this study are consistent with previous findings of processing difficulty by Carminati (2002) when OSP refer to prominent referents.

Iverson (2012) conducted a case study on attrition focused on one adult speaker of L1 Chilean Spanish with L2 Brazilian Portuguese examining different linguistic properties, such as acceptability of NS and OSP and AR, among others. The Chilean control group rated OSP as more felicitous than NS when appearing in TS contexts. They also rated NS as significantly more felicitous than OSP in TC. In an embedded subject interpretation task of AR, the participants were presented both visually and aurally with a slide that contained a written sentence and, in a following slide, a question giving them three choices for an answer (matrix subject, matrix object, both). In the case of ambiguous AR, he found that the Chilean controls consistently interpreted embedded NS as coreferent with the matrix subject while they rejected coreference of embedded OSP with matrix subjects (example 34).

(34)   *La madre preparó una sopa para su hija mientras ella cantaba.*

'The mother prepared a soup for her daughter while she was singing.'

4.2.1.3 Latin American Spanish

Keating, VanPatten and Jegerski (2011), similarly to Alonso-Ovalle et al. (2002), investigated intrasentential anaphora in adult speakers of Latin American Spanish (not strictly monolinguals) as well as HS and L2 learners of Mexican Spanish with English as the dominant language. An offline written questionnaire was used to examine the PAH predictions. The type of sentences was the following: *Juan vio a Carlos mientras ∅/él caminaba en la playa*, 'John saw Charles while (he) was walking on the beach'. The test sentences were 'globally ambiguous', meaning that 'critical items lacked pragmatic, grammatical, or contextual biases in favour of a particular antecedent' (Keating et al. 2011: 207). Each sentence was followed by a

comprehension question regarding the antecedent of the embedded subject. Monolingually raised Spanish speakers showed an antecedent bias only for NS resolution, pointing to a lack of a strong division of labour. There was a strong tendency for Spanish speakers to link NS to the subject of the previous clause, in line with PAH, but the same participants interpreted the OSP as coreferential with the object antecedent around half of the time (as in Alonso-Ovalle et al. 2002).

Similar results were reported in Jegerski, VanPatter and Keating (2011), where monolingually raised Spanish speakers interpreted NS as corefering to subject antecedents, but they showed chance preferences for OSP resolution. The results were consistent with the PAH for Italian only respecting NS interpretation. Resolution of the OSP was different in Spanish compared to Italian since the OSP did not trigger one interpretation over the other. Therefore, Spanish speakers had a weak bias for OSP preferences compared to the strong bias attested for NS antecedence.

4.2.1.4 Argentinian Spanish

Reference was also investigated by Gelormini-Lezama and Almor (2011) in an experimental study on Argentinian (River Plate) Spanish with adult participants. The authors followed Almor's (1999) pragmatic approach to AR, namely the Informational Load Hypothesis, which states that processing cost relates to amount of semantic information that is activated by a referring expression. Anaphor processing reflects a balance between function and processing cost related to mechanisms of verbal working memory. Any anaphoric expression functions as a reactivator of information that is held in working memory in order to achieve a coherent link with the preceding discourse. The authors considered 'repeated names' arguing that these would be always preferred if correct identification of an entity is the only consideration in anaphor processing. However, this was not the case because the capacity of our working memory is limited and the use of (repeated) anaphors, when unneeded, causes superfluous activation of information resulting in processing cost. Thus, repeated names or anaphors in general, unnecessarily used, reflect an activation of redundant information with no discourse function, which increases processing cost.

The rationale of their study was that since the default option for salient antecedents is the NS, repeated names and OSP referring to subject antecedents would incur a higher processing cost. In their reading experiment, they included sequences of two sentences, with a repeated name, an OSP or a NS in subject position in the second sentence as illustrated in (35) (Gelormini-Lezama & Almor 2011: 13).

(35)   *Juan se encontró con María. Juan/Él/∅ la vio triste.*

        'Juan met with María. Juan/He/∅ found her sad.'

Spanish speakers were faster in reading NS in TC contexts (subject antecedent conditions) than repeated names and OSP. The opposite effect was revealed in TS contexts (object antecedent conditions), i.e. reading times were shorter with OSP than with NS. Reading times for the repeated names were a little slower than those for the OSP, while NS in TS were the slowest to process. An example of the object antecedent conditions is (36) (Gelormini-Lezama & Almor 2011: 13).

(36)   *María se encontró con* Juan. *Juan/Él/∅ la vio triste.*

        'María met with Juan. Juan/He/∅ found her sad.'

In TC contexts, NS were easier to process than OSP confirming that more informative forms are redundant. In Spanish 'the default antecedent of the null pronoun is the subject antecedent and forcing the reader to link a null pronoun with an object antecedent requires a re-analysis' (Gelormini-Lezama & Almor 2011: 7). The opposite was true for TS, where the overt forms were found to be faster to process than NS. There was no significant difference between OSP and repeated names in the results, although repeated names were generally read slower than OSP. The NS bias was stronger than the OSP bias given the differences detected in reading times.

The so-called 'repeated name penalty' and the 'overt pronoun penalty' are both conceptualised as penalties for overspecificity in cases of same reference. Their typological distinction is that the former is universal, hence independent of language-specific anaphoric inventories, while the latter is presumably found only in NS languages. For Gelormini-Lezama and Almor (2011: 11), the 'overt pronoun penalty' reveals the 'narrow pragmatic function' of the OSP in NS languages, in which verbal

morphology encodes the grammatical features that in principle suffice for antecedent identification of an anaphoric expression (see also Gelormini-Lezama & Almor 2014). Filiaci et al. (2013) found that when Spanish OSP were associated with syntactically prominent antecedents (subjects), the processing penalty was significantly weaker than in Italian, suggesting that the OSP penalty in Spanish may not be very strong.

4.2.1.5 Mexican Spanish

Shin and Cairns (2012) investigated the development of preferences for NS and OSP in Mexican Spanish. The participants were monolingual Spanish-speaking children, adolescents (aged 6-7, 8-9, 10-11, 12-13, 14-15) and adults. They used a judgment task in which the participants were told brief stories consisting of three sentences, with consecutive grammatical subjects having the same referent (TC) or a different referent (TS). The participants were asked whether the phrase with an OSP or the phrase with a NS was more felicitous for each context. OSP were considered redundant in TC contexts and NS referentially ambiguous in TS contexts.

Adult participants preferred NS in same-reference contexts and OSP in TS contexts, showing a tendency for avoidance of both redundancy and ambiguity. This finding is partially at odds with the previously reported Spanish studies, since OSP were found to be more flexible as regards antecedent preferences, while in Shin and Cairns (2012) this was not the case. The authors claim that preference for OSP in TS seems to be stronger than preference for NS in TC. The evidence provided is that 53% of the adults selected 100% OSP in TS contexts, whereas only 27% of them selected 100% NS in TC contexts. It was argued that preference for OSP in TS represents a propensity to avoid ambiguity while preference for NS in TC manifests a propensity to avoid redundancy. Therefore, avoiding ambiguity seemed to be stronger than avoiding redundancy in adult monolingual native Spanish speakers.

In contrast, children's performance revealed a developmental trend regarding the preference for OSP in TS. The children aged 6-7 did not prefer OSP in TS, while those aged 8-13 did show a preference towards the OSP but not at the adult level. Performance of the 14-15 group was adult-like. In TC contexts, there was no clear

developmental trend since all children groups did not prefer the use of NS. This suggests that in early years there is a tolerance for ambiguous reference, which progressively decreases with age. The children and adolescents' performance indicated that sensitivity to switch-reference contexts develops between ages 8 and 9 and reaches adult levels only at age 14-15. Shin and Cairns (2012) explain children's infelicitous use of pronouns as follows: The children seem to momentarily not consider the interlocutor's point of view when faced with increased processing demands, i.e. complex linguistic tasks or structures. This interpretation is called 'perspective taking' account by the authors (see also Shin & Cairns 2009).

The results in Shin and Cairns (2012) can be compared to those of other studies. In Sorace et al. (2009), Italian monolinguals aged 6-7 also selected significantly more inappropriate OSP in TC than monolingual adults and older children aged 8-10. In Montrul and Sánchez-Walker (2015), Mexican-Spanish monolingual children were significantly different from adults in their higher rate of NS in TS, similarly to Shin and Cairns's (2012) findings. Argyri and Sorace (2007) also showed that Greek monolingual children aged 7;5-9;7 accepted redundant OSP in TC contexts. Comparable patterns have been reported in Papadopoulou et al. (2015) for monolingual Greek children. Namely, younger speakers aged 6-8 did not favour OSP in TS and those aged 6-11 did not reveal a clear resolution pattern for NS, unlike monolingual adults but similarly to Shin and Cairns's (2012) findings for Spanish. In both Greek and Spanish, therefore, resolution routines are inconsistent in young monolingual speakers. Such findings also indicate that the aforementioned aspects of syntax-discourse/pragmatics interface are acquired late (Sorace & Serratrice 2009; Tsimpli 2014; Sorace 2016).

4.2.2 Spanish: Summary

Most studies in Spanish AR reveal that NS antecedent preferences tend to be similar to Greek in adult speakers. NS generally prefer subject antecedents and are mostly used in TC contexts, evidence which partially supports the PAH (Alonso-Ovalle et al. 2002; Montrul 2004a; Montrul & Rodríguez Louro 2006; Gelormini-Lezama & Almor 2011; Keating et al. 2011; Jegerski et al. 2011; Filiaci 2011; Iverson 2012; Shin & Cairns

2012; Filiaci et al. 2013; Bel & García-Alcaraz 2015; Montrul & Sánchez-Walker 2015; Clements & Domínguez 2016; Montrul 2016a). This was not the case in Callahan et al. (2007) on Chilean Spanish, where both NS and OSP could refer back to more prominent antecedents (the subject/topic) at similar rates. Chamorro et al. (2015) also report opposite results compared to most Spanish studies, namely that NS had no bias towards any antecedent manifesting a random or more flexible behaviour.

There seem to be mixed results as for OSP antecedent preferences and contexts of use in Spanish. While some studies show that OSP do not strongly select an antecedent (Alonso-Ovalle et al. 2002; Callahan et al. 2007; Keating et al. 2011; Jegerski et al. 2011; Filiaci 2011; Filiaci et al. 2013), other studies show that OSP do have a clear antecedent bias towards the less prominent antecedent or that OSP are clearly preferred in TS contexts (Montrul 2004a; Montrul & Rodríguez Louro 2006; Gelormini-Lezama & Almor 2011; Iverson 2012; Shin & Cairns 2012; Chamorro et al. 2015; Montrul & Sánchez-Walker 2015; Clements & Domínguez 2016; Montrul 2016a). There seems to be a difference in the production and interpretation modality regarding OSP distribution in Spanish. In interpretation, OSP tend to be more flexible while in production OSP are mostly used in TS contexts.

With regard to the strength of bias in favour of one antecedent over the other in AR, most Spanish studies report a stronger bias for the NS pronoun and a weaker bias for the OSP in antecedent assignment (Alonso-Ovalle et al. 2002; Callahan et al. 2007; Keating et al. 2011; Jegerski et al. 2011; Filiaci 2011; Gelormini-Lezama & Almor 2011; Filiaci et al. 2013; Bel & García-Alcaraz 2015; Clements & Domínguez 2016). On the other hand, Shin and Cairns (2012) and Chamorro et al. (2015) report a contrastive finding, namely that the OSP bias is stronger than the NS in adult performance.

Evidence for Spanish subject distribution and AR has not been conclusive so far. Different studies have examined different varieties of Spanish. There is variation in OSP use across varieties (Mayol 2012) and possibly in related discourse/pragmatic constraints. Thus, it cannot be taken for granted that results from one Spanish study can be generalised about the Spanish language. Difference in data elicitation methodology is also a key contributing factor to the varying results.

4.3 Discussion

A comparison between Greek and Spanish related research as a whole would suggest that in general Greek studies are relatively consistent in their findings, despite methodological heterogeneity, as compared to Spanish studies. NS favour assignment with subject antecedents in most Greek and Spanish studies. By contrast, OSP appear to differ in the two languages, with the Spanish OSP evidently being weaker than the Greek OSP. It seems to be the case that the OSP in Spanish does not always comply with the discourse feature of salience/prominence to the same degree as in Greek. Production vs interpretation modalities in research should be also taken into account as they seem to reveal different patterns in participants' performance.

In sum, the Greek OSP has a strong bias towards TS contexts since it triggers consistent interpretations in AR. Although NS in Greek are generally used in TC contexts by default referring more often to subject antecedents, they seem to be unstable compared to their overt counterparts in AR. In Spanish, preferences for NS referring to subject antecedents are generally stronger than those found for OSP. Spanish OSP are less stable than NS, while in Greek NS are less stable than OSP. However, by and large NS' coreference to subject antecedents seems to be a commonality in the two languages.

From all the aforementioned findings it can be inferred that more similarities are expected in NS than in OSP between Greek and Spanish. When compared, the two languages should exhibit stronger differences in the distribution and interpretation of OSP and fewer differences in the distribution and interpretation of NS. The next question is to examine whether Greek-Spanish bilinguals use the same subject assignment strategy as monolinguals in production and interpretation given the differences in the ability to integrate syntactic and pragmatic/contextual information, as seen in Chapter 3. The following chapter focuses on the research questions and the methodology of the present study.

Part II

5. The study

5.1 Introduction

The current study explores subject distribution in Greek in contact with Spanish in Chile in adult populations. In order to factually establish whether and to what extent the two languages are equivalent in subject distribution, the first goal is to directly compare monolingual performance between Greek and Chilean Spanish (henceforth Spanish) in both production and interpretation using the same methodology. The second goal is to examine Greek in contact with Spanish in order to discover differences from monolingual Greek in subject distribution. It is hypothesised that bilingual Greek performance should be different due to language contact, which may result in crosslinguistic influence and/or different processing of subject properties located at the interface. Comparisons are drawn between the different bilingual and monolingual groups' production and interpretation of subjects in a methodologically consistent way. Comparing bilingual groups sheds light into the potential sources of differences between monolingual and bilingual Greek.

Therefore, the aim of the present study is twofold:

(a)   It seeks to directly compare Greek and Spanish monolinguals in order to discern differences in the scope of third-person null and overt subjects between the two languages. Previous evidence suggests crosslinguistic differences, which however have not been adequately investigated using the same methodology.

(b)   It attempts to directly compare Greek data from monolinguals and Greek-Spanish bilinguals in order to identify potential differences in the distribution of third-person null and overt subjects in Greek in a situation of language contact with Spanish and explain its potential sources by comparing different types of bilinguals.

5.1.1 Outline of the predictions

As regards the monolingual speakers, it is assumed that subject distribution is guided by similar discourse/pragmatics conditions in Greek and Spanish. However, these conditions are not necessarily identical in the two languages. The scope of OSP is expected to be wider in Spanish and narrower in Greek. Monolinguals may occasionally be redundant in use and/or interpretation of overt subjects, since referential biases in the contexts of interest are not categorical. In production, NS in TS could obtain without resulting in pragmatic inappropriateness, while in AR, due to lack of context, NS may be more ambiguous.

The default assumption would be that bilinguals of two NS languages should not differ from monolinguals since no obvious differences in subject distribution are observable at first sight. The prediction, however, is that bilingual groups would differ from Greek monolinguals in their Greek performance because subject distribution is a domain shown to be unstable and vulnerable to variation under crosslinguistic influence. In production, bilinguals may misuse third-person overt subjects and NS, i.e. may be redundant and also under-explicit in different contexts. In interpretation of ambiguous anaphora, they are expected to be more flexible or inconsistent in antecedent choices compared to monolinguals. Particularly, a relaxed markedness of the Greek OSP is expected in the bilingual performance as has been typically found in previous research on NS languages following assumptions stemming from the IH. The bilingualism effect is expected to be stronger in HS and L2ers than in immigrants.

The predictions guiding the two studies of the present dissertation are formulated in detail in §6.1 and §7.1 for the production and interpretation tasks respectively.

5.1.2 Sources of potential differences

If bilinguals are different from monolinguals, constraints due to interface conditions enhanced by crosslinguistic differences (if found) could be considered as plausible sources of differences in the first place. Keeping interface conditions and processing cost as a potential constant in performance in the less used/less dominant language, other important variables in the context of the current study include Age at the time

of testing (Age) and Proficiency. The study takes a lifespan approach considering adult speakers of all ages, as shown in §5.2.6. Therefore, Age is a crucial variable in the statistical analyses (regressions), which is expected to explain some of the findings (see Kaltsa et al. 2015). As will be shown in §6.1 and §7.1, a further research question in all cases concerns the association of participants' age with their linguistic performance. For HS and L2ers, proficiency as well as literacy in Greek are factors taken into consideration as potentially playing a role in their linguistic behaviour.

The bilingual groups represented different types of bilingualism with different sociolinguistic characteristics. For immigrants, the length of residence in Chile (LoR) was strongly correlated with Age (see §5.2.5.2). For HS, the number of years of bilingualism (YoB) coincided with their age, since age of onset of bilingualism (AoB) was birth. For the L2ers, these variables involved the age at which they started learning Greek. The length of temporal experience of being bilingual indicates the degree of language contact intensity. The variables concerning speakers' language acquisition/learning patterns and bilingual experience were factors that could explain aspects of linguistic behaviour and were considered in the discussion of the findings.

5.2 Methodology

5.2.1 Methodological approach

The present study focuses on use and interpretation of the most common referring expressions for subject reference, i.e. NS, OSP and LS, examining the interaction between syntax and discourse/pragmatics. Data from Greek and Chilean Spanish monolinguals as well as from different types of Greek-Spanish bilinguals in their Greek performance were elicited and compared employing the same methods.

Pinto (2014: 82) contends that 'all instances of language development or language transition (including all the many cases of bilingualism in which I include L1 attrition) may show a discrepancy between comprehension and production'. The present study examines both elicited production and interpretation of subjects. Such an approach allows insights into the strategies employed by monolingual and bilingual speakers in

oral production and aural comprehension discerning potential discrepancies in the two modalities (Argyri & Sorace 2007). In addition, language use may not accurately reflect linguistic competence, due to performance limitations (e.g. hesitations, false starts). Spoken language comprehension, on the other hand, involves a receptive task, which may unveil the underlying grammar as mirrored in speakers' intuitions. Thus, a combination of two methodological paradigms, which are usually employed independently, examines Greek and Spanish both separately and in contact in the context of bilingualism. Such an approach allows a fuller picture of the speakers' competence and performance (see also Rothman 2007b).

The data were obtained from semi-spontaneous speech elicited from oral narratives and interpretation of ambiguous anaphora given in aural form to the participants. The use of the written modality in the target language was discarded because of the minority situation of Greek in the bilingual context of this study (see Sitaridou & Kaltsa 2014; Austin, Blume & Sánchez 2015; Montrul 2016b). In order to capture a good picture of speakers' linguistic performance, both a natural communication task (narratives) and a linguistic manipulation task (AR test) were used. The former entailed free production of contextualised structures chosen by the participants in a spontaneous way following storylines of picture sequences, which instigates usage of grammar in a natural way targeting the implicit, unmonitored linguistic knowledge. The latter focused on decontextualised stimuli, presented aurally to the participants in order to elicit their intuitions on AR strategies in oral speech. Both tasks were considered to largely involve the most commonly used linguistic structures in day-to-day oral linguistic communication. However, while language production and interpretation are often multimodal insofar as using and/or observing body gestures is involved, the present methods were based exclusively on oral and auditory skills.

The focus on third-person is based on the fact that within the person feature there is a split between first/second and third person (Carminati 2005; Lozano 2009; Lubbers Quesada & Blackwell 2009; Pinto 2012; Schmitz & Scherger 2017). As seen in §2.3.6.1, the Greek third-person pronoun *aftos* has a special status in being identical in form with the demonstrative. Third person has been found to be more vulnerable at the interface than the other two persons, since overproduction surfaces especially with

third-person singular animate pronouns but not with first/second person pronouns (Lozano 2009). The feature coming next in importance is number, with [singular] being the default, thereby being more 'basic' than plural (Adger 2003; Carminati 2005; Lozano 2009; Mayol 2012). Third person and singular number are the defaults in person and number systems respectively (Ackema & Neeleman 2017a). Thus, the focus is on singular number in interpretation and the majority of utterances in narratives. In production, plural number also emerged, albeit to a much lesser degree than singular. As to gender, in the interpretation task only feminine and masculine referents were included because of lack of neuter nouns in Spanish.

5.2.2 Procedure for data collection

As Scontras, Fuchs and Polinsky (2015: 16) observe, first 'we must understand the complexities of the multilingual experience before we can analyse its exponence in language users'. The methodological approach adopted for the present research incorporated features of the social network approach, i.e. studying language in the situational context of speech community (Milroy 1987). An ethnographically-oriented procedure for the bilingual data collection was followed. Data collection was carried out in situ in Chile and in Greece in four fieldwork trips and it was conducted in two rounds corresponding to two trips in each country (see Table 5.1). The first round dealt with elicitation of production data and the second one with interpretation data.

Table 5.1. Fieldworks for data collection in Chile and in Greece in chronological order

Field trips	Country	Period	Duration	Main action
1st	Chile	Nov. 2015 - Feb. 2016	3 months	Interviews, narratives
2nd	Greece	Mar. - Apr. 2016	10 days	Narratives
3rd	Greece	Dec. 2016	10 days	Anaphora resolution
4th	Chile	Jan. - Feb. 2017	$1^{1/2}$ month	Anaphora resolution

An asset for the transatlantic data collection process was my fluency in Greek (native) and Spanish (near-native) as well as my familiarity with the Chilean sociocultural context due to previous residence in Chile. My status as a researcher in the data-collection process was that of both an insider and an outsider (Milroy 1987). Each participant received personalised attention in a spirit of trust. Absolute confidentiality

and anonymity were guaranteed in order to avoid ethical issues. A meaningful understanding of the speakers was facilitated by approching them and interacting with them in their own social environment, which contributed to ecological validity.

The participants in the Chilean context were recruited through personal contacts, word of mouth and snowball sampling. The study reached the wider community of Greek-Spanish bilinguals in Chile and, given the relatively small number of speakers (see §1.2), the data obtained are considered to be largely representative. Another aspect related to the present study is that Greek is a highly regarded European language within the Chilean society. The positive stance towards the language and its speakers, along with the small size of the Greek-speaking population, allowed a well-disposed response from the participants smoothing the way for an effective approach to their sociolinguistic profiles and linguistic behaviour.

Data collection was completed during two field trips in Chile, whose duration was three months and six weeks respectively. Before and during the first fieldtrip, contact was established with a number of (potential) participants. All pertinent institutions (e.g. Greek communities and the Greek Embassy) were informed about the study and requested to spread the word on participants' recruitment. A detailed sociolinguistic profile was collected for each bilingual participant through a comprehensive (though not exhaustive) background questionnaire and an oral biographical interview (§5.2.3).

Prior to the first fieldtrip in Chile, many of the participants who had informally consented to participate in the study were contacted by email. All relevant information was sent to them, together with a consent form and the background questionnaire (Appendix A, B). They were requested to read and sign the consent form and to complete the questionnaire, to send those back by email and to notify their availability for the interview. All texts, information and instructions were given in the language which was more convenient to the participant. In the case of older participants who were not familiar with electronic means of communication, contact was established through telephone calls, with them or their relatives. In the latter cases, the forms were filled out during the meeting, in which all information was presented to these participants orally.

The first round of data collection in Chile was accomplished with one-to-one oral interviews. Detailed sociolinguistic information and the narratives using the picture-story task were elicited and saved in digital recordings (Tascam DR05 V2 recorder). In the next fieldwork trip, the second round of data collection focused on the interpretation task, which was conducted in person with roughly the same participants. In all cases, the interviews took place in naturalistic settings and were individual to give each participant complete attention and to avoid potential pressure from third parties. In some occasions, the participants were also seen (and recorded) together with relatives in customary interaction before or after the main interview.

The meetings were held at a prearranged time in a location proposed by the participants. In most cases, people suggested their homes and less frequently so their workplace. Occasionally, the meetings took place at quiet cafes and a couple of times at a park. The meetings with the participants lasted from 30 minutes to 4 hours with most meetings lasting around 1.5 hour. The interviews of the first round were longer than those of the second one due to the biographical interview, for which there was not a restricted time frame. The participants keen to talk were welcomed to do so for as long as they wished (see §5.2.3.2). In addition, the participants were volunteers; they only received small gifts from Greece or Cambridge, sweets or flowers when I was invited to their home and treats in cases of interviews in cafes.

The monolingual participants of Greek and Spanish were recruited through personal contacts in Athens and Santiago. They signed the consent form and answered a short questionnaire on their general and linguistic background to assure that they were monolingually raised and had minor/no knowledge of other languages. They performed the two tasks in the first and second round of data collection respectively, without participating in a biographical interview. The Spanish-speaking monolinguals were tested during the aforementioned field trips in Chile. The Greek monolinguals were tested during two additional field trips in Greece. The time lag between the two rounds for data collection in both countries (see Table 5.1) was due to practical reasons. Further details on data elicitation as well as on data manipulation after data collection are given in following sections.

5.2.3 Sociolinguistic profiling of bilingual speakers

Individual variability is the hallmark of HS populations because of the wide variation in their linguistic experience and contexts of language acquisition and use (Montrul 2016b). The same holds for immigrants and L2ers. The sociolinguistic situations in which bilinguals speak the minority language must be adequately understood. Detailed sociolinguistic information of the bilingual participants of the present study was collected via two methods: a questionnaire and a biographical interview.

5.2.3.1 Background questionnaire

A detailed questionnaire seeking to capture basic aspects of the participants' biographic and sociolinguistic profile was collected before the interviews, either via email or at the actual meeting. The questionnaire was two-page long and divided into two parts. The first part focused on the participants' general background enquiring on principal ethnographic information, such as age, place of birth, educational level, age at immigration (if relevant) and frequency of visits to Greece. The second part of the questionnaire sought to capture the participants' linguistic background regarding language practices across lifespan. Information was requested on the participants' first and dominant language, language(s) spoken by the parents, estimates of language ability, language exposure and use as well as knowledge of other languages (see Appendix B). The questionnaire was answered in written form by the participants who were familiarised with this sort of practice and in oral form by the rest of participants (e.g. the older ones) with my assistance.

There was a Greek and a Spanish version of the questionnaire. The participants chose the language that was more convenient for them. Most HS and L2ers preferred the Spanish version, while most immigrants preferred the Greek version. The information gathered from the background questionnaire offered insights into the extralinguistic variables that could influence the participants' linguistic knowledge at testing. The questionnaire responses were also useful in providing a primary idea of each speaker's profile before the interview and thus shaping the content of the biographical interviews on topics therein explored more thoroughly.

5.2.3.2 Biographical interview

The qualitative oral interviews followed a semi-structured design focusing on the participants' sociolinguistic details, intended to be given in a life-story narrative in the context of an informal conversation (see Pauwels 2016). Conducted entirely in Greek, in addition to capturing the participants' sociolinguistic profiles, part of their interview was also used to evaluate HS and L2ers' Greek proficiency (§5.2.7). There was no set time limit for the participants to speak, thus the length varied considerably among the interviewees depending on factors such as inhibition, eloquence, availability of time and level of proficiency (lower proficiency speakers were generally less talkative). They were encouraged to maintain a conversation for at least 10 minutes, which was the minimum length of a biographical interview, while the maximum was around 150 minutes, with most interviews lasting between 15-25 minutes. The interviews were digitally recorded and orthographically transcribed.

Most questions were open-ended and followed a flexible interview protocol. Topics included family histories or personal histories regarding migration (if relevant), their linguistic and educational background, Greek and Spanish speaking and learning experience and practices as well as frequency of language exposure. The interviews began with a warming-up stage with a short ice-breaking chat. The actual interviews started with a question on the participants' place of birth and upbringing. The conversations moved forward with tracking their family history and the rest of the themes depending on the speakers' profile and considering their preferable topics for further elaboration. There was no strict control over the content since I followed up leads, but the main topics across the recordings were generally comparable. The qualitative data gathered through the biographical interviews, although not quantifiable, were illustrative in terms of sociolinguistic information from the participants' perspectives and in their own words, aspects which would not have emerged from a mere questionnaire. The data provided an insightful understanding of the speakers and the reliability of sociolinguistic information was assured to the greatest extent possible. After the biographical interview, in the same session the participants were requested to perform the oral production task.

5.2.4 Tasks

Both instruments employed for eliciting production and interpretation data were oral so that speakers were not confronted with the written language. As has been noted in the literature, a large number of HS are unfamiliar with formal spoken and written registers and they hardly know how to read or write (e.g. Montrul 2016b). In the case of Greek, the difference in the writing system (i.e. orthography) is an extra complication taken into account. The tasks were selected upon the basic criterion of involving informal spoken registers, familiar events, uncomplicated vocabulary and common linguistic structures. Additionally, the tasks were considered appropriate to elicit speakers' intuitions with respect to the linguistic structures in question. In this regard, there were neither hints nor opportunities for metalinguistic thinking in the performance of the tasks. Thus, data elicitation followed a natural procedure targeting speakers' subconscious, implicit linguistic knowledge in both production and interpretation. All data were collected in a quiet space and all participants were naive to the specific objectives of the tasks. There was no time pressure, aside from the AR task in which the participants were prompted to respond as soon as they heard each sentence and question.

5.2.4.1 Study 1: Production (Story-telling task)

Production data were collected using the two picture story description tasks by Hickmann (2003), the *Horse Story* and the *Cat Story,* consisting of sequences of successive pictures that form two stories with animal characters (Appendix C). The story-telling task elicited oral narratives containing semi-spontaneous use of third-person referential subjects in an ordered discourse structure (story generation). In narratives, speakers are free to select the referring expressions of their choice and all constructions are contextualised in series of temporally-connected events. This method aims to evoke rules or principles that govern language use 'in the wild', i.e. how speakers naturally form propositions to produce coherent pieces of contextualised discourse. Such data tend to be representative of authentic language use. The method has been extensively employed in related research on subject expression and reference (e.g. Silva-Corvalán 1994; Hendriks 2003; Montrul 2004a,

2016a; Tsimpli & Sorace 2006; Belletti et al. 2007; Arnold et al. 2009; Iverson 2012; Leclercq & Lenart 2013; Pinto 2013, 2014; Tsimpli, Andreou, Agathopoulou & Masoura 2014; Hendriks, Koster & Hoeks 2014; Andreou 2015; Montrul & Sánchez-Walker 2015).

The picture-sequences consisted of black and white drawings presented to participants in paper in a pre-established order. The *Horse Story* consisted of five pictures and the *Cat Story* of six pictures (Appendix C). The stories differ in the status of their animate referents. In the *Horse Story*, the main protagonist is a horse, which is portrayed in all pictures, it is mostly agentive and it appears to be in full view and bigger than the other two secondary animal characters (cow, bird). This story works well for eliciting expression of TC contexts (see Hendriks 2003). In the *Cat Story*, there is no clear difference in the status of the characters. The story starts by depicting a bird family with three baby birds, which are not agentive, and the mother bird, which is seen leaving. The cat then seems to play a central agentive role, which is however overshadowed by the appearance of the dog. Hickmann (2003: 184) observes that the *Cat Story* 'requires a more complex temporal organisation, since it shows main events that overlap or occur in close temporal proximity. [...] the cat arrives as the mother bird flies away, the dog arrives as the cat climbs up the tree, the mother bird returns as the dog pulls down the cat'. Consequently, this story narration seems to be a good instrument to elicit natural expressions of TS contexts.

It should be noted that the animal characters participating in the stories have not a strictly determined gender in Greek since, depending on the noun chosen, the referent can have masculine, feminine or neuter gender. For instance, the noun 'cat' in Greek is usually feminine (*i gata*), but can be also masculine (*o gatos*) or neuter (*to gati*). Similarly, the noun 'dog' is typically masculine (*o skilos*) or neuter (*to skili*). An additional related aspect, thus, is the grammatical consistency in gender assignment in referential expressions, such as OSP or clitics, especially by bilingual speakers. Wrong or inconsistent gender cues may mislead referent interpretation. This is important when it comes to track reference of antecedents in narratives. As shown in §6.3.2.3, ambiguity of NS referents is sometimes related to the referents' gender.

The data collection method took into account certain methodological considerations, in line with Hickmann (2003). The participants were presented with the two picture sequences and they were instructed to narrate a story with each sequence as accurately as possible. Before starting the narration, they were encouraged to ask me about any unknown words that they would need, since vocabulary knowledge was not a target domain in the present research. In this way, pressure to retrieve vocabulary, which could affect oral performance, was minimised.

As Hickmann (2003: 182) notes, 'mutual knowledge is [...] clearly a most central aspect of speech situations, determining the uses of a variety of linguistic devices'. Although the pictures were visually available to me, the participants were crucially instructed to pretend to tell the stories to an imaginary listener who could not see the pictures and did not know the stories (see also Hendriks et al. 2014). It was thus assumed that there was no situation of shared knowledge between the interlocutors, which would affect the choice of referring expressions (Sorace 2004).

The speakers were not allowed more than a few minutes to plan their speech; hence the data were of an informal and semi-spontaneous nature. An advantage of eliciting production with pictorial stimuli is that it helps to avoid memory difficulties since the to-be-narrated content remains constantly visible to the participants. Since the narrators do not rely on short-term memory, the process of coherent speech production is facilitated. Nevertheless, pictures provide static frames which could invite deictic uses instead of discourse-internal ones (Hickmann 2003), the latter being the target of this study. The participants were thus prompted to (try to) use a narrative rather than a descriptive mode in their production.

In bilinguals, the narrative task was performed right after the biographical interview. For both stories, the average duration of recordings was 3-4 minutes. All utterances were transcribed using standard orthographic transcription. The data were manually coded and annotated in an Excel database, controlling for category of subject, type of predicate, person, number and information structure, among others (see §5.2.8). Each utterance containing a verb (each clause) was the basic unit of analysis (see Polinsky

2008; Shin 2016). Non-referential subjects and formulaic chunks, among other structures, were excluded from the analysis (see details in §5.2.8.1).

The production task aimed at eliciting narratives which, because of the events of the stories, lend themselves particularly well to the study of subject expression in the contexts of interest. In producing prompted narratives, choices with respect to use of subjects, such as type (null/overt), category (lexical/pronominal), definiteness, position and information structure were made by the speakers in a relatively natural context. The focus was, thus, on narrative microstructure, which means concentrating on linguistic features, such as morphosyntax and referential devices, used in the construction of coherent discourse (see Gagarina, Klop, Kunnari, Tantele, Valimaa, Baleiuniene et al. 2015; Gagarina, Klop, Tsimpli & Walters 2016).

The picture-story narrative tasks were performed orally first in Greek and then in Spanish on the same occasion by bilingual participants. Despite the known problems arising from conducting the same task in two languages on the same day, it was deemed inevitable due to practical reasons related to the difficulty to reschedule a second meeting with the participants. However, the quality of the Spanish production data was not optimal since in the Spanish version, which followed the Greek one, the speakers tended to (more or less) translate their Greek narrative into Spanish. For this reason, the Spanish version of the narratives was not included in the present study.

5.2.4.2 Study 2: Interpretation (Anaphora resolution task)

The interpretation data were obtained using an experimentally based method, which was an off-line, self-paced listening task in the form of an oral comprehension questionnaire, modified from Mastropavlou, Katsiperi, Fotiadou, Fleva, Peristeri, Tsimpli (2014) on Greek monolingual adults. Their study was an online self-paced listening and picture-matching task, which took into account the PAH (Carminati 2002) and, additionally, the role of definiteness in AR (see §4.1.1.1). The main differences between the original study and the present one, in methodological terms, have to do with the particular characteristics of the bilingual population which participated in this study and are the following:

(a)   In the present study, the task was not picture-matching as in the original study. The participants only listened to recorded sentences and then answered orally a comprehension question (included in the recording) for each sentence. The decision to exclude the pictures was partly based on the fact that the definiteness was manipulated in the object antecedent. Picture stimuli could affect participants' antecedent choices, since the image of a definite and an indefinite object antecedent (e.g. 'the nurse/a nurse') was exactly the same. Moreover, despite the task being untimed, it was crucial that the participants would answer the question immediately after listening to each recording. Displaying pictures would possibly complicate more the process of responding. Visual stimuli do provide memory cues, but may confuse participants causing an additional processing cost in deciphering them. This point is particularly relevant in cases of elderly participants, a fair amount of which participated in this study. Therefore, in order to avoid visual errors and possibly random answers, no use of pictures was made for this task. It was also considered that hearing and interpreting oral speech is a natural and habitual process of daily life, which does not necessarily demand visual cues.

(b)   The AR task did not offer options of possible answers/antecedents for the participants to choose, contrary to the original study, in which three potential referents were displayed in images (subject, object, 'other'). Instead, the participants were instructed to answer orally to the comprehension question giving their first intuition. The first data were collected in Greece and were initially used on a pilot basis, which was finally included in the main study with no modifications. In the initial data collection stage, none of the participants ever responded that the agent of the verb in question was a third referent (absent from the linguistic context). The experimental items with the respective questions, therefore, led the listeners to select one of the antecedents comprised in the matrix clause. Additionally, in Mastropavlou et al. (2014) the percentages of preference for the 'other' alternative were very low. Thus, no different options and no 'other' referent were given to the participants, since these would not be particularly beneficial to the realisation of the task.

(c)   Manipulation of the matrix subject definiteness was not included in the present study because constructions with indefinite subjects are considered to be relatively infrequent according to general principles of given-before-new ordering of

information within sentences. Moreover, indefinite subjects are generally used for newly introduced referents. In the production study, such presentational/existential sentences were excluded from the main analysis, since the focus was on TC and TS contexts, which involve referents' maintenance and reintroduction. Thus, only experimental items with definite subjects in the matrix clause were considered.

(d)   The length of the oral questionnaire was shorter than the one used in the original study. This decision was taken considering the fact that it is easy for the participants to get tired of answering lengthy questionnaires, especially in a language which is not dominant for them. In particular, in interpreting ambiguous anaphora structures, if the process takes too long for them, it may happen that they start giving random answers in order to finish as soon as possible. It was, thus, regarded convenient to reduce the size of the questionnaire so that the participants could stay focused, thereby securing the best possible reliability in their responses.

(e)   The interpretation task was an offline rather than an online test, in the sense that the participants' reaction times were not recorded. This is because of the wide age range of speakers taking part in the study (16-87), with a number of them being elderly. It was thus deemed that measuring reaction times would be ineffective since the older participants would evidently show a slower processing speed than the younger ones. Moreover, this would possibly not reflect actual reaction times but rather speed of hand movements, irrelevant to the present research. The possibility of computer use on the part of the participants was also rejected since some older people dislike or feel intimidated by computers due to unfamiliarity. Therefore, the task focused only on the participants' antecedent preferences in resolving null and overt pronouns in ambiguous anaphora structures elicited through their oral responses, with no measure of reading or reaction times.

The AR task tested interpretation of non-biased referentially ambiguous intra-sentential forward anaphora, involving biclausal discourse contexts. Forward anaphora is generally preferred to backward anaphora (cataphora), thereby being more frequent across languages as the 'unmarked' word order (Blackwell 2003; Iraola 2015). The test sentences presented a sequence of two events in a subordinating discourse structure. There were two third-person singular referents in a matrix clause

in the canonical word order (SVO), a subject and an object, matched in gender. The matrix clause was followed by an adverbial (temporal) clause consisting of either a NS or an unstressed OSP matched in gender and number with the two matrix antecedents. In the NS condition, an adverb was placed in the position of the subject to maintain the same number of segments across conditions. Half of the test sentences included feminine referents and the other half masculine referents. The verb of the embedded clause was always in past imperfect. Two variables were manipulated: the anaphoric subject in the embedded clause and also the definiteness of the object in the matrix clause. The constructions were fully ambiguous because the referents in the matrix clause were equally prominent and antecedent of the embedded pronoun could either refer to the subject or the object of the matrix clause. An example of an experimental item is (37), in which the matrix object is indefinite and the embedded subject is OSP. All the experimental sentences and the filler items are displayed in Appendix D.

(37)   a. *O ðiefthindis xeretuse enan jatro otan aftos evjene apo to asanser.*
         Pjos evjene apo to asanser?
      b. *El director saludaba a un doctor cuando él salía del ascensor.*
         ¿Quién salía del ascensor?
         'The director was greeting a doctor when he was exiting the lift.
          Who was exiting from the lift?'

The conditions thus were four, with the matrix subject being definite in all cases:

(i)   Definite matrix object - Null embedded subject   (DDN)
(ii)   Indefinite matrix object - Null embedded subject   (DIN)
(iii)   Definite matrix object - Overt embedded subject   (DDO)
(iv)   Indefinite matrix object - Overt embedded subject   (DIO)

The test sentences were 16 (8 with a NS, 8 with an OSP) and were presented in a randomised order, with 1:1 ratio to the fillers and two practice items at the beginning of the task (see Appendix D). Equivalent versions of the oral questionnaire in Greek and Spanish were created. The Spanish version was used for the Spanish monolinguals only. The experimental sentences and questions were recorded by a

male native speaker of Greek and Chilean Spanish respectively, who produced the items naturally with clear voice and flattened prosody (neutral intonation). The recorded sentences were saved in a tablet (Samsung Galaxy Note 10.1) and were presented auditorily to the speakers requesting them to use earphones (their own or provided by me).

In order to reduce metalinguistic awareness, the participants were prompted to answer as quickly as they could. After listening to each item, the participants had to answer orally who was the doer of the action of second clause. Given that the referents in the matrix clause were of the same gender (masculine or feminine), the participants had to choose between two competing antecedents in subject or object position. In exceptional cases, they could listen to an item for a second time. The participants were recorded and an entry of their responses was made in written form. Subsequently, the data were verified and inserted into a database in Excel.

Contrary to the production task, which involved contextualised discourse units expressed by the participants, the interpretation task concerned decontextualised discourse units presented to the participants. The experimental sentences were considered to be completely ambiguous without providing any cues (e.g. world knowledge, grammatical cues) that could help to resolve NS or OSP, thus the pronouns could not be identified through any mechanisms other than the use of the anaphor in the particular narrow linguistic context in which it appeared.

As seen in §2.4.3, listeners deal with ambiguity by having preferences as captured by the PAH (Carminati 2002). Subjects are frequently omitted in non-contrastive non-emphatic contexts since referents can be usually recovered directly from verbal morphology. Thus, subject omission occurs mostly in TC contexts. Change of referent (TS) is typically realised by an overt subject, which is LS in referent introduction and OSP or LS in referent reintroduction. In interpreting ambiguous anaphora in the present study, NS and OSP involved referent continuity or reintroduction. The ambiguity of the sentences served to determine whether pronouns have clearly identifiable biases. The purpose of the task was to find to what extent the PAH is

operative in non-biased ambiguous intra-sentential contexts in Greek and in Spanish as well as in Greek in contact with Spanish, using the same methodology.

Narratives allow a fine-grained exploration of the contextual factors regulating AR. The experimental design of the interpretation task provided a testing ground for the PAH in isolated sentences, which lack context and are thus expected to guide the participants to spontaneously apply (or not) referential biases. However, two points should be taken into account regarding production and interpretation respectively:

(a) Oral production may be affected by performance limitations, such as articulatory or memory constrains, resulting in unintentional speech errors which do not truly reflect speakers' competence. This is why aural interpretation is important in providing a means of better assessing linguistic knowledge (see Rothman 2007b).

(b) In resolving ambiguous anaphora, in lack of context the listeners may silently fill in contexts in the otherwise decontextualised sentences. As Altmann and Kamide (2007: 502) state, 'sentence comprehension is not a passive process that projects an articulated world onto some inner mental screen. Instead, it is a process that results in active behaviours directed towards the contents of the concurrent world'. The participants may invent a context in their mind and this extralinguistic information ultimately affects their preferred interpretation in resolving ambiguous reference.

Another point to consider, as noted in literature, is that experimental research on AR has foremost considered PAH-like contexts, which constitute only 'one narrow instance of the numerous and complex set of discourse structure variables that could influence pronominal reference' (Jegerski et al. 2011: 503). On this ground, use of subjects in narratives may illuminate AR from different angles regarding expression (and thus expected interpretation) of particular referential forms.

5.2.5 Participants

Five main groups of participants took part in the study, namely (a) two monolingual groups consisting of Greek and Spanish speakers; and (b) three bilingual groups consisting of immigrants (potential L1 attriters), heritage speakers (HS) and L2 speakers (L2ers). There was a difference in the group composition between first and second round of data elicitation, more prominent in the monolingual groups and slighter in the case of bilinguals. Thus, production and interpretation data were obtained from somewhat different groups of participants. This was due to the time lag between the two rounds of data collection. Consequently, some speakers participated only in the production task and some other speakers participated only in the interpretation task. In the case of bilinguals, both tasks were performed by mostly the same participants.

Concise information about the monolingual and bilingual groups can be seen in Tables 5.2 and 5.3 as well as in Appendix E. Differences in participants' age between Study 1 (production) and Study 2 (interpretation) exist due to two reasons: (a) interpretation data were collected after (more or less) a year from the eliciting production data (thus, the participants who took part in both tasks were one year older at the second-round testing); and (b) there were (slight) differences in group composition. The same holds for LoR in immigrants, i.e. in AR those who participated in both tasks had one year extra in their LoR. The overall age range was 16-87, involving adult individuals with no (or not obvious or known) pathological problems related to language. Crucially, the older participants did not suffer from any significant age-related cognitive decline.

Table 5.2. Groups of speakers participating in Study 1 (production task)

Groups of speakers		N	Age	Years in Chile	N of Clauses
Monolinguals:	**Greek**	20 (13 fem)	29-70 mean: 45.4 (SD: 12.7)		982
	Spanish	20 (7 fem)	28-77 mean: 45.2 (SD: 13.7)		890
Bilinguals:	**Immigrants (all)**	35 (21 fem)	22-86 mean: 51.1 (SD: 22.3)	1-66 mean: 24.6 (SD: 25.2)	1574
	Younger Immigrants	23 (14 fem)	22-52 mean: 35.7 (SD: 6.4)	1-20 mean: 7.4 (SD: 5.6)	1034
	Older Immigrants	12 (7 fem)	74-86 mean: 80.6 (SD: 3.5)	29-66 mean: 57.6 (SD: 10.2)	540
	Heritage Speakers	21 (7 fem)	16-67 mean: 52 (SD: 13.7)		678
	L2 Speakers	20 (8 fem)	24-79 mean: 44.4 (SD: 17.3)		715
Total		116			4839

Table 5.3. Groups of speakers participating in Study 2 (interpretation task)

Groups of speakers		N	Age	Years in Chile
Monolinguals:	**Greek**	20 (9 fem)	16-80 mean: 48.1 (SD: 19.8)	
	Spanish	20 (10 fem)	20-85 mean: 47.8 (SD: 19.3)	
Bilinguals:	**Immigrants (all)**	34 (21 fem)	22-87 mean: 54 (SD: 22.4)	1-68 mean: 27.7 (SD: 25.9)
	Younger Immigrants	21 (13 fem)	22-53 mean: 37.3 (SD: 6.7)	1-21 mean: 8.5 (SD: 6.1)
	Older Immigrants	13 (8 fem)	72-87 mean: 81 (SD: 4.4)	30-68 mean: 58.7 (SD: 9.9)
	Heritage Speakers	23 (8 fem)	17-85 mean: 54.7 (SD: 15)	
	L2 Speakers	18 (7 fem)	24-72 mean: 41.7 (SD: 16.1)	
Total		115		

5.2.5.1 Monolingual speakers

Greek and Spanish monolingual speakers (N=40) were recruited in Greece and Chile and were 20 adult individuals per group. In the Greek group, 11 speakers participated in both tasks while 9 speakers participated in only one task, either the production or the comprehension task. In the Chilean group, 8 participants performed both tasks while 12 participants completed only one task. The criteria for selecting adult monolingual native speakers were: (a) to be Greek or Chilean national; (b) to reside in Athens or in Santiago; (c) to have been monolingually raised in Greek or in Spanish; and (d) to have low/hardly/no proficiency in other languages. All Spanish-speaking monolinguals were speakers of the Chilean variety only. All monolinguals had attended at least twelve years of education. Tertiary-level education was completed by 14 and 11 Greek monolinguals in the production and interpretation task respectively as well as by 13 and 10 Spanish monolinguals in the respective tasks. An attempt was made to incorporate Greek and Chilean adult participants of a similar age span in order to have homogeneous monolinguals groups, which could also match bilinguals in terms of age to the extent possible. The lowest age limit was set at 16 years, since the youngest HS in the study was 16 years old at the first-round testing. Tables 5.4 and 5.5 display the age range of monolinguals in the two tasks.

Table 5.4. Age of Spanish and Greek monolingual speakers in oral production task

Study 1: Age of monolingual speakers in years					
Language	N	Mean	SD	Min	Max
Spanish	20	45.20	13.72	28	77
Greek	20	45.45	12.65	29	70

Table 5.5. Age of Spanish and Greek monolingual speakers in interpretation task

Study 2: Age of monolingual speakers in years					
Language	N	Mean	SD	Min	Max
Spanish	20	47.80	19.30	20	85
Greek	20	48.15	19.82	16	80

A Wilcoxon-Mann-Whitney test was performed and showed no significant difference in the variable Age between the two monolingual groups for each set of data, i.e. production (z=-0.217, p=0.828) and interpretation (z=-0.095, p=0.924).

5.2.5.2 Bilingual speakers

A total of 84 bilingual speakers participated in both data collection rounds. The selection criteria for recruiting bilingual participants were set as follows: The participants should (a) be bilinguals (in the broad sense of the term) of Greek and Spanish, i.e. they should have (at least some) fluency in Greek and in Spanish; (b) be adults of any age; and (c) live permanently in Chile. The following speakers were a priori excluded from the study: (i) speakers of Cypriot Greek; (ii) those who had very low proficiency in Greek or in Spanish; (iii) those who had lived in several countries and/or were (raised as) multilinguals; (iv) first-generation of Greek immigrants who had been living in Chile for few years reporting no actual immersion in Spanish and regular use of English; (v) those who were not residing permanently in Chile; and (vi) linguists or language-related professionals. The large majority of bilinguals resided in Santiago. Interviews were also collected from participants living in Valparaíso (two speakers), Viña del Mar (three speakers), Antofagasta (two speakers), Puerto Montt (one speaker) and Punta Arenas (four speakers) (see Figure 1.1).

The bilinguals were divided in three groups according to basic differences in their biographical and bilingual traits. Common denominators include (a) Greek-Spanish bilingualism in Chile, a context geographically distant from Greece; (b) sharing the same language contact situation, with Spanish being the dominant societal language and Greek a minority language, thereby sharing the potential effects of crosslinguistic influence; and (c) speaking Greek as a heritage or minority language due to Greek origin or other strong cultural or family connection. Differentiating factors include age of onset of bilingualism (AoB) (simultaneous/early or sequential/late bilingualism), order of acquisition, language of schooling, literacy skills, frequency of exposure, input quality and language acquisition setting (naturalistic or/and instructed). Some speakers reported visits to Greece on a regular basis[11], whereas most of them (especially older immigrants, HS and L2ers) reported no such opportunity. This is a contributing factor in estimating quantity and quality of input.

---

[11] This piece of information is indicative of participants' socioeconomic status (SES), since travelling from Chile to Greece is a costly activity. However, SES is not taken as a variable of interest in this study.

First-generation immigrants (IMM)

This group encompasses first generation of Greek immigrants in Chile, who were born and schooled in Greece, raised monolingually and residing in the host country for periods of varying length. In line with Schmid (2011: 7), first-generation immigrants or L1 attriters are 'speakers for whom the onset of attrition (i.e. the moment of migration) took place after the onset of puberty'. The group of immigrants was composed by late sequential bilinguals since they arrived at a linguistically mature age in Chile and learned L2 Spanish in early adulthood (except for one case outlined below). The immigrants constitute the parental generation with respect to HS. A number of 35 speakers participated in the production task and 34 in the interpretation task. Among them, 29 participants performed both tasks, 6 participants performed only the production task and 5 participants performed only the interpretation task. All these speakers (N=40) reported having advanced or near-native proficiency in Spanish, which was the language predominantly used in every-day situations. Table 5.6 shows information on the immigrants' age.

Table 5.6. Age of immigrants in production and interpretation tasks

	Immigrants: Age in years				
	N	**Mean**	**SD**	**Min**	**Max**
Study 1	35	51.11	22.29	22	86
Study 2	34	54.02	22.38	22	87

The minimum length of residence (LoR) in Chile was one year and the maximum was 66 years in the production data and 68 in the interpretation data (Table 5.7). Although speakers with few years in Chile are less likely to present L1 attrition signs than those with many years of residence in the host country, all first-generation immigrants are sometimes called attriters (see Seliger 1991; Schmid & Köpke 2017a, 2017b). There is no clarity as to when L1 attrition may appear as this depends on many factors. More than 10 years (Montrul 2008) or at least 6 years (Tsimpli et al. 2004) of intense exposure to the majority language may give rise to attrition. The immigrant speakers with few years of LoR crucially reported being fully immersed in Spanish (living with Chilean families and using Spanish at work and daily life).

Table 5.7. Length of residence in Chile in production and interpretation tasks

	Immigrants: LoR in Chile (N of years)				
	N	**Mean**	**SD**	**Min**	**Max**
Study 1	35	24.62	25.23	1	66
Study 2	34	27.70	25.89	1	68

Age and LoR were strongly correlated (narratives: Spearman's rho=0.748, p<0.001; AR: Spearman's rho=0.792, p<0.001) (Figures 5.1 and 5.2).
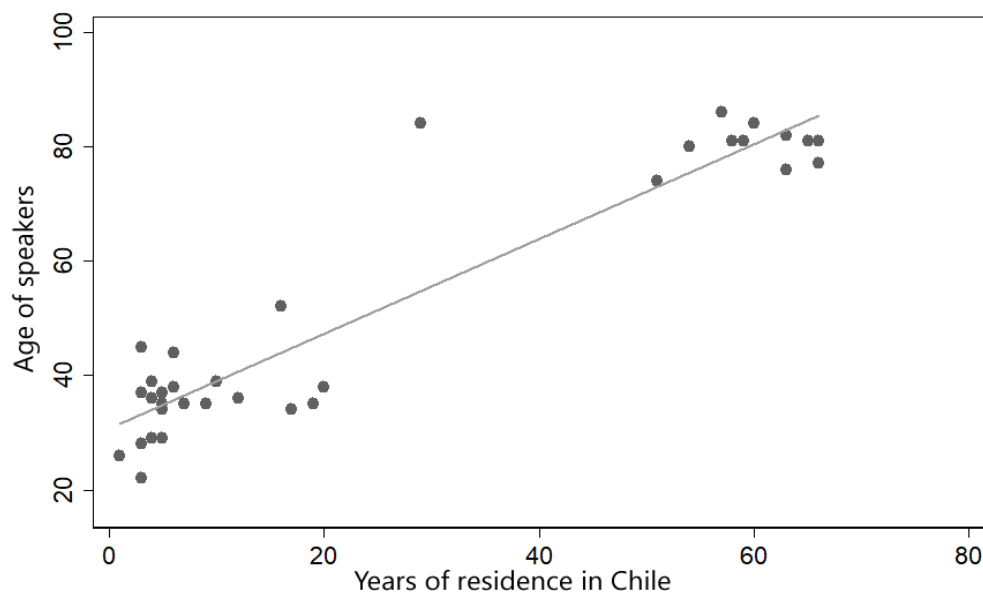
Figure 5.1. Scatter plot: correlation between age and length of residence in narratives
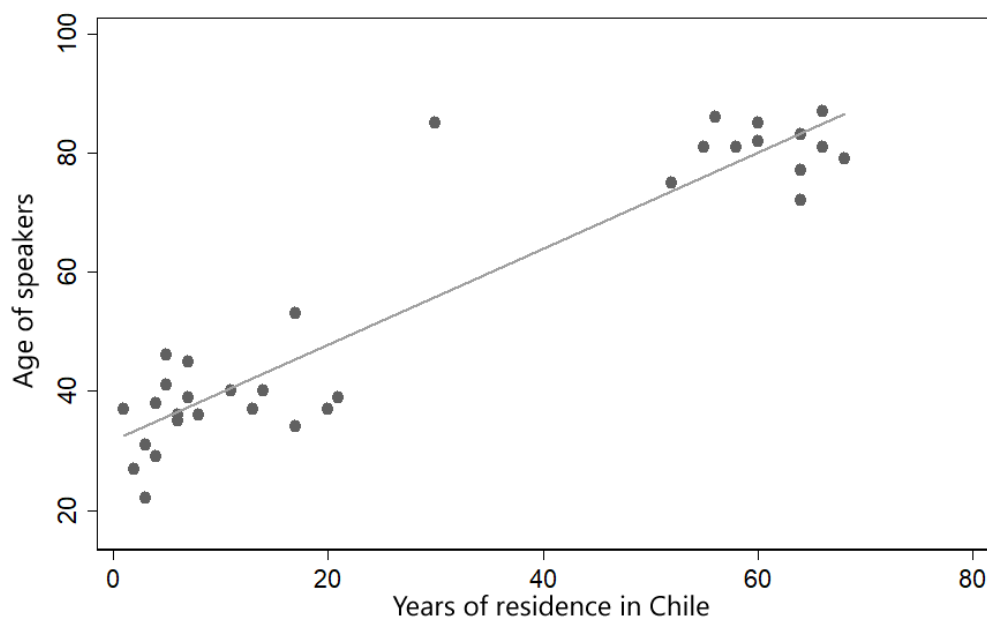
Figure 5.2. Scatter plot: correlation between age and length of residence in AR

Thus, greater LoR implied more advanced age of participants. The immigrants were divided into two age groups (younger, older). The younger group consisted of speakers of 22-52 and 22-53 years old in the narratives and AR respectively, while the age range for the older group was 74-86 and 72-87 in the two tasks respectively, as shown in Tables 5.8 and 5.9 (see Kaltsa et al. 2015).

Table 5.8. Age of younger and older immigrants in oral production task

Study 1: Age of immigrants in years					
Group	N	Mean	SD	Min	Max
Younger immigrants	23	35.73	6.41	22	52
Older immigrants	12	80.58	3.47	74	86

Table 5.9. Age of younger and older immigrants in interpretation task

Study 2: Age of immigrants in years					
Group	N	Mean	SD	Min	Max
Younger immigrants	21	37.28	6.70	22	53
Older immigrants	13	81.07	4.42	72	87

All the older immigrants had received education in Greek and knew how to read and write in Greek and Spanish. A total of 5 out of 14 older participants had completed only primary-level education, 3 out of 14 had attended secondary-level education and 6 out of 14 had received tertiary-level education. Their AoB coincided with the age at migration. From the 26 younger immigrants who participated in both tasks, 23 had completed tertiary-level education and 3 had received only secondary-level education. The younger immigrants' AoB did not always coincided with age at migration since some of them had learned L2 Spanish before resettling to Chile.

For all the younger immigrants, Spanish was in fact an L3 or L4 if a chronological order of attending L2 courses is taken into account. Thus, all the younger immigrants reported that they had studied (at least) English before learning Spanish. Spanish is regarded as an L2 here not only for convenience purposes but also considering the following fact: Spanish was the dominant societal language, in which at least an advanced level of command had been attained, while the other language(s) were not regularly used by these speakers. All the older immigrants learned Spanish as an L2 strictly speaking.

There was one speaker who arrived in Chile before puberty, namely at the age of 8 (S162). The participant performed only the interpretation task. She was included in the group of immigrants rather than in that of HS pondering the following factors: (a) her age at migration was relatively late in childhood; (b) both her parents were Greek; (c) she was (by far) the youngest of three siblings in the family (her brother and sister being young adults at migration), hence the input at home was entirely monolingual provided by end-state mature native speakers of Greek; and (d) the family's high SES allowed an increased frequency of trips to Greece. Due to these factors, it was deemed that she had more commonalities with first-generation immigrants than with HS (although this may be debated, since most of her formal schooling experience was in the host country and in Spanish). The participant in question would be a case of what some scholars call 1.5 generation (Suarez 2007) in the context of heritage language maintenance and loss.

Heritage Speakers (HS)

The HS were children of first- or second-generation immigrants. They had been exposed to Greek at home from birth and spoke (or used to speak) Greek as their family language. The group thus comprised simultaneous bilinguals who were dominant in Spanish (unbalanced). Similarly to monolinguals, the HS acquired Greek naturalistically through exposure to primarily aural input in conversational every-day contexts, but in terms of quantity and quality, the input which they received was different from that received by monolinguals.

A number of 21 and 23 HS participated in the production and interpretation tasks respectively, with 21 participants responding both tasks and 2 of them participating only in the interpretation task. All HS were born in the host country, except for two cases, which involved immigrant children who arrived in Chile from Greece at the age of 4 and 6 years respectively (S23 & S139) (1.5 generation). In both these cases, the participants had one Chilean parent and one Greek parent, so they were also exposed to both languages from birth. Within this group, 18 participants were second-generation speakers, while 5 of them were third-generation speakers. Table 5.10 provides information on the age of HS.

Table 5.10. Age of HS in production and interpretation tasks

Heritage Speakers: Age in years					
	N	Mean	SD	Min	Max
Study 1	21	52.00	13.71	16	67
Study 2	23	54.69	15.01	17	85

Oral proficiency in Greek ranged from basic to near-native, with most of the HS being assessed as having an intermediate level (see §5.2.7). Additionally, 11 HS reported that they had studied Greek at some point (self-taught or formally instructed) as heritage language learners (Polinsky & Kagan 2007; Montrul 2016b), but in most cases literacy skills were hardly or not fully mastered. Only one HS had attained advanced reading and writing skills in Greek (S25). As for education and literacy, all HS were schooled in Spanish in Chile, with 6 out of 23 having received secondary-level education and 17 out of 23 having completed tertiary-level education.

Many of the HS reported that they had undergone attrition in their L1 Greek (in the sense of substantial reduction of language input and use, resulting in partial forgetting), whose trigger was marked by one or two determining events: (a) the onset of formal education in the dominant/majority language around the age of 6; and (b) their Greek-speaking parent(s)' death at some point in the HS' later life in the case of older participants.

As regards the parental background, in terms of Kupisch (2013), 14 out of 18 second-generation speakers were 100% heritage in the sense that both their parents represented the heritage language being first-generation immigrants. Moreover, two of them were 50% heritage because only one parent was Greek (in both cases the father). Additionally, two participants could be evaluated as being 75% heritage since one parent was a first-generation Greek immigrant and the other was a Greek HS. As previously mentioned, five participants were third-generation speakers, with four of them having Greek grandparents from one side (maternal or paternal) and in only one such case from both sides. These speakers were thus children of HS. They had acquired the heritage language at home primarily receiving input from their parents and grandparents (i.e. from HS and immigrants respectively).

Unlike monolinguals and immigrants, HS received quantitatively less and qualitatively different input in Greek. In terms of input quality, most HS had not developed formal literacy skills in the heritage language, thus they were not familiar with reading and writing, neither with formal registers (except S25). Some of the relatively old HS were not familiar with vocabulary which is considered as core at the present-time language. For example, three HS (100% heritage, aged 54-59) of intermediate-advanced fluency did not understand the word *tenies* ('films') in Greek. It is worth noting that these three speakers had never been to Greece.

In the HS group, variability aside, the common denominator was that the speakers acquired Greek at home as a heritage language (or as an L1 which became a heritage language soon in life in cases of immigrant children), i.e. through early naturalistic exposure in a bilingual context. HS had 'more temporal experience at being bilingual' in comparison to both immigrants and L2ers (Pascual y Cabo & Rothman 2012: 5).

L2 Speakers (L2ers)

The group of L2ers included late sequential bilinguals with L1 Spanish who had learned L2 Greek in adulthood. Their difference from typical L2 learners was their cultural or family connection with Greek. These speakers could fall into the category of 'broadly defined heritage speakers', for whom 'the heritage language is equivalent to a second language in terms of linguistic competence, and as a second language, it typically begins in the classroom, in adulthood' (Polinsky & Kagan 2007: 369).

The narratives task was performed by 20 L2ers and the AR task by 18, with 17 of them taking part in both tasks, 3 completing only the production task and one of them only the interpretation task. As regards education, 3 out of 21 L2ers had finished secondary-level education only, while 18 out of 21 had attended tertiary-level education. For 8 out of 21, Greek was an L3 considering the chronological order in which they had learned the language. Greek is referred to as L2 for all speakers for convenience purposes but also considering recency and frequency of use. Although some speakers had attended L2 courses in other languages (e.g. English) before learning Greek, they reported no habitual use of the other language and more

opportunities for exposure to Greek (or pursuit thereof). Table 5.11 gives information on the age of the L2ers.

Table 5.11. Age of L2ers in production and interpretation tasks

	L2 Speakers: Age in years				
	N	Mean	SD	Min	Max
Study 1	20	44.40	17.32	24	79
Study 2	18	41.72	16.09	24	72

A number of 12 of these speakers were third-generation of Greeks, grandchildren of original immigrants, who learned the heritage language as adults in an instructed setting. Some of them (N=4) had been exposed to a low or minimal degree to the heritage language as receptive bilinguals or overhearers in childhood. None of them spoke the language before adulthood, according to their reports. Additionally, 4 L2ers were not descendants of Greek immigrants but had a Greek consort (3 cases involving marriage with children). Moreover, 5 L2ers were not of Greek origin neither had they family bonds with Greeks, but for family reasons they had migrated to Greece for a fair amount of time (4-25 years)[12] and they had returned to Chile.

The L2ers varied as to their incentives to learn L2 Greek in adult life. The common characteristic shared by all members was a strong cultural, personal or family connection to the language, which was in some way associated with migration. Their motivation for learning Greek 'would thus contrast significantly with that of the typical foreign language student' (Valdés 2001: 2). From this perspective, Greek could be regarded as a heritage language for these speakers in the broadest sense of the term, although they substantially differed from typical HS. The main difference was that Greek was an L2 learned in an instructed setting, while for HS it was an L1 naturalistically acquired in a bilingual context at home.

Major differences between L2ers and immigrants, both late sequential bilinguals, were the following: (a) order of language acquisition, i.e. L1 Greek-L2 Spanish for the immigrants and L1 Spanish-L2 Greek for the L2ers; (b) most of the older immigrants learned L2 Spanish through immersion, while all the L2ers learned Greek through

---

[12] These speakers could be regarded as L2 attriters.

instruction, at least in the first place; (c) many of the older immigrants had to learn L2 Spanish out of necessity (circumstantial bilingualism), while the L2ers learned Greek primarily guided by choice and self-motivation and not due to need (elective bilingualism); (d) the shared dominant language of society was Spanish, i.e. the L1 for the L2ers, but the L2 for the immigrants. It could be assumed that the immigrants were generally more balanced in their two languages compared to L2ers. This is because late L2 onset of bilingualism often coincides with L1 dominance; thus, each bilingual group could be generally regarded as being (more) dominant in their L1. L2ers' L1 was the majority language, hence they were assumed to manifest not as much tendency to use Greek. For the immigrants, the minority language was still relatively dominant (especially for the younger ones), at least more dominant than for the L2ers (see Montrul 2008). As regards performance in the minority language, Sorace (2016: 9) observes that 'in L2 speakers, the unwanted language is the (still) dominant L1, which requires more resources to be inhibited; in L1 attriters, in contrast, the unwanted language is the (less dominant) L2, which requires fewer resources to be inhibited'. Furthermore, the older the immigrants, the longer their LoR, hence more dominant in Spanish due to greater intensity of language contact.

The L2ers' oral proficiency in Greek ranged from basic to near-native, with most participants being intermediate and near-native speakers (see §5.2.7). According to self-reports, all had attended formal classes to learn L2 Greek, thus literacy skills were mastered to a certain extent in all cases. Formal instruction in classroom settings clearly leads to enhanced metalinguistic awareness as opposed to naturalistic context immersion in which HS acquire the heritage language. Evidently, the participants who had lived in Greece had an abundance of opportunities for naturalistic exposure and interaction in an immersion context compared to the other speakers of the L2 group. The same proportionally holds for the participants whose close family members (e.g. consort, children) are speakers of Greek and share the same household in Chile.

5.2.6 Age

The age of the participants at the time of testing is depicted in Tables 5.12 and 5.13 concerning all groups in the narratives and the AR task respectively.

Table 5.12. Age at testing in all groups of speakers in oral production task

	Study 1: Age of speakers					
	Group	**N**	**Mean**	**SD**	**Min**	**Max**
Monolinguals	Spanish	20	45.20	13.72	28	77
	Greek	20	45.45	12.65	29	70
Bilinguals	IMM	35	51.11	22.29	22	86
	HS	21	52.00	13.71	16	67
	L2	20	44.40	17.32	24	79
	Total	116	48.12	17.24	16	86

Table 5.13. Age at testing in all groups of speakers in interpretation task

	Study 2: Age of speakers					
	Group	**N**	**Mean**	**SD**	**Min**	**Max**
Monolinguals	Spanish	20	47.80	19.30	20	85
	Greek	20	48.15	19.82	16	80
Bilinguals	IMM	34	54.02	22.38	22	87
	HS	23	54.69	15.01	17	85
	L2	18	41.72	16.09	24	72
	Total	115	50.13	19.39	16	87

A one-way ANOVA was performed to analyse the association between Group of speakers and Age. The results showed that the model was not significant [narratives: $F_{(4,111)}=1.03$, $p=0.397$; AR: $F_{(4,110)}=1.67$, $p=0.162$], indicating that the different groups of speakers were not significantly different in the distribution of age. The age range within each group of speakers was wide, therefore Age at the time of testing was a crucial variable considered in the statistical regressions (see §5.2.6).

When the association between Age and AoB was analysed, there was a significant correlation between these variables in both datasets (narratives: Spearman's rho=-0.116, $p<0.001$; AR: Spearman's rho=-0.154, $p<0.05$). Therefore, the variable AoB could not be used as an independent variable along with Age.

5.2.7 Proficiency

According to Grosjean's (2008: 23) complementarity principle, 'bilinguals usually acquire and use their languages for different purposes, in different domains of life, with different people. Different aspects of life require different languages'. The level of fluency attained in language skills is determined by the need for these skills, which is often domain specific in the case of bilingualism. Reading and writing skills are not developed in a language unless necessary. Thus, proficiency can display a broad continuum in the heritage language in different contexts (Polinsky & Kagan 2007; Benmamoun et al. 2013; Montrul 2016b; Rothman, Tsimpli & Pascual y Cabo 2016).

The degree of oral proficiency of HS and L2ers ranged from basic ability to fully fluent and native-like, with most speakers being at intermediate level and above. The different proficiency degrees were related to speakers' particular language acquisition trajectory and experience. Sociolinguistic determinants are intrinsically related to language proficiency within the variability of bilingual practice (see Montrul 2016b). Proficiency 'has an impact on language processing and thus it affects bilinguals' performance on […] tasks that involve informants' language processing mechanisms' (Treffers-Daller 2009: 74).

The HS and the L2ers (N=41) were assessed on their oral performance only since most of the former reported low, minimal or no competence in reading and writing skills in the heritage language. They had a (more or less) stable level of bilingualism even when the ultimate level of proficiency attained in Greek was relatively low because they were not attending classes at the time of testing. The proficiency levels were four: basic, intermediate, advanced and near-native.

Standardised language proficiency tests involving written language could not be used in the current study, therefore an alternative method of assessment was utilised focusing directly on oral productive skills and indirectly on receptive skills. The proficiency in these two groups of speakers was assessed based on the three pieces of information mentioned below.

5.2.7.1 Self-reports

In the background questionnaire, the participants stated their estimated level of proficiency in Greek (basic, intermediate, advanced, near-native, native) as well as how often they heard, spoke, read and wrote the language. These answers served as a first clue to speakers' proficiency, which was then crosschecked and complemented in the biographical interview. It is well known that in self-rating speakers often do not evaluate themselves accurately, i.e. they may overrate or underrate their language abilities (Grosjean 2008; Austin et al. 2015). Taking into account the proficiency level as indicated in the questionnaire, I subsequently assessed the speakers' oral production during the biographical interview in line with the below-stated criteria.

5.2.7.2 Evaluation criteria for oral production assessment

During the interview, the participants' productive skills were rated according to the criteria applied in assessing candidates sitting the official exam for the Certificate of Attainment in Greek as established by the Centre for the Greek Language (see Council of Europe 2001, 2018; Centre for the Greek language 2012). The main evaluation grid is holistic. Taking as a point of reference the speakers' estimated level as indicated in the questionnaire, the following specific yardsticks of oral proficiency were used to measure it: (a) grammatical accuracy; (b) pragmatic appropriateness; (c) pronunciation and intonation; (d) vocabulary; and (e) effectiveness in communication. Each criterion was given a score from 1 to 5, depending on level of reference, with 5 indicating the best performance and thus adding up to 25 (100%) as the highest possible score. If the overall performance was graded with a minimum of 15 (60%), the participant could be evaluated as being at the level of reference. If the grade was below 15, the evaluation indicated a lower level. If it was 25 (100%), it was then determined whether the participants' performance indicated the level of reference or a higher level. The evaluation was made online during the interview or soon thereafter and it was effective given my experience in assessing candidates for Greek as an L2 in the aforementioned formal examination. This allowed an overall impression of the speakers' proficiency. Part of the participants' speech was further scrutinised after being transcribed to get an objective diagnostic of their proficiency.

5.2.7.3 Grammaticality index

Grammatical knowledge as expressed in spontaneous speech was measured by deviations from monolingual native baseline norms. The participants' oral proficiency was assessed using a grammaticality index on the basis of the (un)grammaticality of 50 clauses composing a random part of each speaker's transcribed oral biographical interview. The interviews contained naturalistic data elicited without rigorous control over the content. The clauses used for the assessment were those forming the first part of the participants' biographical interview excluding the very first phrases and any formulaic chunks. The grammatical and ungrammatical clauses were counted focusing on the expression of grammatical phenomena such as agreement, case marking or aspectual distinctions. The percentage of the ratio of error free units was used to establish the grammaticality index (Table 5.14; Appendix E).

Table 5.14. Grammaticality Index indicating speakers' oral proficiency

Grammaticality Index - Proficiency	
% of grammatical clauses	Proficiency
62-72	Basic
73-83	Intermediate
84-94	Advanced
95-100	Near-native

Scrutinizing grammatical accuracy reflected the speakers' proficiency in oral production since monolingual native speakers rarely commit grammatical errors (usually these are only slips of the tongue). The proficiency scale for both groups of speakers (HS and L2) was established as shown in Table 5.14 and the numbers of speakers per level is depicted in Figure 5.3 and Table 5.15.

It can be seen that there is no 'native' level of performance, although all HS are regarded as native speakers of their two languages, as mentioned in Chapter 1. This is because no speaker in the HS and L2 groups revealed a monolingual native performance considering their overall Greek performance. That said, there were cases of HS and L2ers with near-native level of proficiency in Greek.
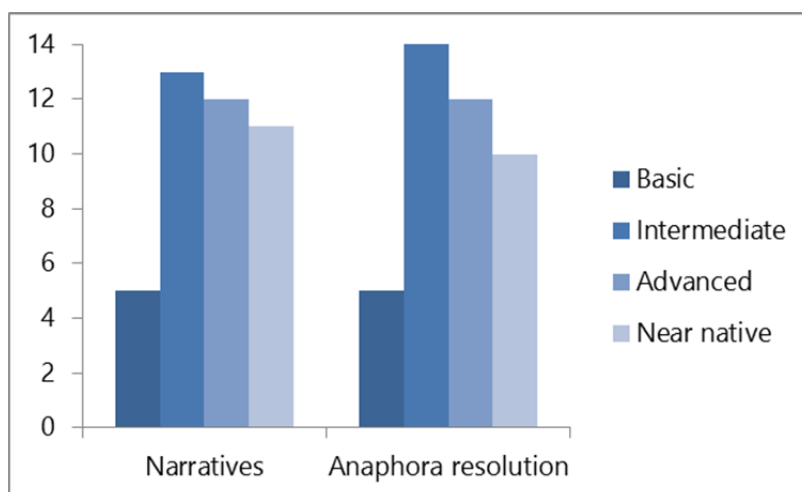
Figure 5.3. N of speakers per proficiency level per task

Table 5.15. N of speakers per proficiency level per group

	Narratives			AR		
	HS	**L2**	**Total**	**HS**	**L2**	**Total**
Basic	1	4	5	3	2	5
Intermediate	8	5	13	8	6	14
Advanced	7	5	12	7	5	12
Near-native	5	6	11	5	5	10
Total	21	20	41	23	18	41

5.2.7.4 Observations

Lexical diversity was not taken into account in an exhaustive way as in the case of grammaticality to assess oral proficiency. This is because the two groups involved different types of bilinguals, who had acquired the language through largely different input and modality, a fact which possibly affects vocabulary range and lexical richness. The HS were simultaneous bilinguals, i.e. native bilinguals, they had acquired the language primarily via aural input and they were familiar with the vernacular spoken within particular domains and mainly in the context of home. On the other hand, L2ers were late sequential bilinguals, who had learned the language, at least in the first instance, through instruction, i.e. via both written and aural input. The latter speakers may be characterised as 'hyperliterate' (Montrul 2014: 293). The distinction between naturalistic vs instructed bilingualism inevitably entails different types and ranges of vocabulary as used in different registers and social contexts. As mentioned above, some HS acquired the variety of Greek spoken by their parents at the moment

of migration, which took place several decades ago. Some words of contemporary basic everyday vocabulary were not part of these varieties (e.g. *tenies* 'films'). These words, however, are normally included in the L2 teaching and learning materials for any modern language. Hence, measuring lexical diversity through e.g. verbal fluency tests would not necessarily indicate speakers' proficiency in a comparable way for the two groups.

The participants' speaking fluency in terms of speech rate was not measured through words per minute counting, since the recordings occasionally did not accurately represent the participants' real speaking pace. In at least two cases of HS it was observed that the participants were fluent and talkative in heritage Greek before the actual interview, but at the onset of recording their fluency was significantly decreased due to self-consciousness. Moreover, (semi-)spontaneous speech is often interrupted by pauses, repetitions and false starts, especially in a weak language such as a heritage language, affecting speech rate (Polinsky & Kagan 2007).

Proficiency in L2 Spanish in the immigrants was specified considering only their self-reports (questionnaire, interview). For the other bilingual groups (HS, L2), Spanish was the/an L1, regarded as primary and dominant at the time of testing.

5.2.8 Procedure for data analysis

5.2.8.1 Linguistic variables

The oral narratives from 116 speakers were transcribed following standard orthographic transcription including pauses. The data were coded and annotated in a database in Excel, which was converted into a database in Stata (version 11.2). The database, after being finalised and cleaned, comprised 4,839 clauses coming from 232 (=116*2) narratives. The annotation included information on the linguistic variables shown in Table 5.16. Only the linguistic variables in bold were taken into account in the final quantitative analyses of the production data.

Table 5.16. Linguistic variables of the annotation of narratives

Clause & Verb	Subject
· **Type of Clause (matrix or embedded)**	· Type of Subject (overt, null)
· Dependency (if embedded)	· **Category of Subject (NS, LS, OSP)**
· **Type of embedded (adverbial, relative, complement: indicative, subjunctive, infinitive)**	· Category of Subject if overt (personal pronoun, demonstrative pronoun, LS, proper name, demonstrative noun phrase, relative pronoun or relative complementiser)
· Coordination	· **Person**
· Type of verb (existential, unaccusative, unergative, transitive)	· **Number**
	· **Mismatch in SV agreement**
	· Adjacency
· Finiteness	· Position
· **Verb Person**	· Definiteness
· **Verb Number**	· Animacy
· **Ambiguous verb morphology (Spanish)**	· **Discourse Value (TS, TC, Focus)**
	· **Ambiguity**

The following subject structures were excluded from the analysis: non-referential subjects or impersonal verbs (e.g. *prepi* 'must'), direct speech, fixed expressions or fillers (e.g. *pos na pume* 'how to say'), first person (e.g. *vlepume* 'we see'), nominalization of clauses (e.g. *to na traviksi tin ura tis gatas* 'the pulling of the cat's tail'), proverbs (e.g. *mana ine mono mia* 'there is only one mother'), formulaic phrases (e.g. *ke zisane afti kala ki emis kalitera* 'they lived happily ever after'), codeswitching, verb phrase ellipsis, false starts, incomplete sentences and any unclear utterances.

Subject-headed relative clauses were not considered in the narratives analyses because of their particular dependence on the head of the noun phrase. Verbs of relative clauses cannot be used with a NS or an overt subject like in other types of clauses, thereby often being excluded from such kind of analysis (e.g. Dimitriadis 1996; Montrul & Rodríguez Louro 2006; Shin 2012, 2016). The contents of the production task triggered mostly third-person animate subjects in singular. Instances of the other persons were ruled out from the analyses. Plural number and inanimate referents, although rare, were included. The contexts in narratives considered for analyses included only those of TC and TS. Infinitives in Spanish, whether adjuncts or complements, were regarded as clauses (see Torrego 1998; Zagona 2002).

A qualitative approach to the linguistic data was also conducted for each category of subject (LS, OSP, NS) in each context (TC, TS). A detailed examination of all the chains and relations established between matrix and embedded clauses was beyond the scope of the present study.

As for the AR data, the participants' responses were also coded and annotated in an Excel database, which was converted into a Stata database comprising 1,840 responses coming from 115 speakers. The linguistic variables of interest in this case were (a) type of embedded subject (NS, OSP); (b) definiteness of the matrix object (definite, indefinite); and (c) antecedent preferences (AP).

5.2.8.2 Sociolinguistic variables

The language-external variables included in both Stata datasets were Sex, Group of speakers, Age, AoB, LoR (if immigrant) and Proficiency (if HS or L2). For monolinguals, only Sex, Group of speakers and Age were added in the corresponding datasets. The variables that were finally used in the analyses were Group of speakers, Age and Proficiency (Table 5.17). Sex was included only for descriptive purposes. AoB and LoR were strongly correlated with Age, thus only Age was included in the final statistical analyses since it was the only crucial language-external variable common to all groups of speakers of this study (see Kaltsa et al. 2015; Schmitz & Scherger 2017).

Table 5.17. Sociolinguistic variables used in the analyses

Sociolinguistic variables
• Group of speakers: Spanish monolinguals, Greek monolinguals, IMM, HS, L2   • Age at testing   • Proficiency (if HS or L2): basic, intermediate, advanced, near-native

5.2.8.3 Baseline

The ideal baseline against which to compare the linguistic behaviour of bilingual HS has encouraged discussion in the field, with the language of input generally being the most reasonable option. According to Polinsky and Kagan (2007), 'the baseline language for a heritage speaker is the language that he or she was exposed to as a child'. As Benmamoun et al. (2013: 134) explicitly state regarding HS, 'crucially, the baseline language is not the monolingual variety of that language but the language spoken by first-generation immigrants' (see also Polinsky 2018). Schmid (2011) and Grosjean (2008) have also argued that using the monolingual norm as a yardstick to evaluate bilinguals more generally may be questionable. However, the comparison between monolinguals and bilinguals is not only difficult to ignore but it may be particularly useful for answering specific research questions (Montrul 2016b). Such a comparison may disentangle behaviours which are due bilingualism effects from behaviours stemming from cognitive biases which are not language specific. The second research question of the present study involves a comparison between monolinguals and bilinguals in order to pinpoint potentially existing differences and their sources. In order to better understand the sources of potential differences, a comparison between the bilingual groups becomes also relevant.

Comparisons were first drawn between monolingual Spanish and Greek data in order to establish potential differences between the two languages. The focus thereafter was on the Greek performance of the bilingual groups. The bilingual performance was compared against that of Greek monolinguals. The performance of HS and L2 groups was also compared against that of immigrants. The latter speakers are the main source of linguistic input for HS (see e.g. Zombolou 2011; Kaltsa et al. 2015; Montrul 2016b) and often for L2ers in this context. Crucially, HS and L2ers have different AoB and language exposure. Comparisons were also drawn between HS and L2ers since both groups share the characteristic of speaking Greek as their weak language.

5.2.9 Statistical analyses

Statistical analyses were conducted using Stata (version 11.2). The level of statistical significance was established at $p<0.05$. In order to analyse the association between categorical variables (all linguistic variables), Pearson Chi-square tests were performed. Fisher's exact tests were used when the number of data in at least one cell was less than 5. Tests of proportions were used to analyse proportions within a group (e.g. OSP vs NS in a specific group of speakers). The aforementioned statistical analyses were performed in order to obtain an overview of the association between two categorical variables. A Wilcoxon-Mann-Whitney test, which is a non-parametric test, was performed in the description of the data in two cases involving the variable Age due to the fact that the distribution of Age was not normal (see §5.2.5.1 and §6.3.2.3). The data were also submitted to logistic regression analyses, as explained in the following section, which allow more complex analyses including the effect of language-external factors, such as age, in the linguistic performance of the speakers.

5.2.9.1 Logistic regressions

In order to analyse different models, multinomial logistic regressions were conducted in the narratives database and binomial logistic regressions were conducted in the AR database. Both multinomial and binomial logistic regressions belong to the General Linear Models family, which relate a dependent variable with a set of independent variables. While in other linear models, such as ANOVA or ordinary least square regression, the dependent variables are continuous, multinomial and binomial logistic regression models use categorical variables as dependent variables. The dependent variable in the narratives database was categorical with three categories (NS, LS, OSP), thus multinomial logistic regression was the most suitable model to analyse the relationship between the independent sociolinguistic variables and the use of a subject form in production. As for the AR dataset, the dependent variable was categorical representing the preference of the speakers with regard to subject reference (preceding subject or object). Binomial logistic regression was the most suitable model for analysing the relationship between the independent sociolinguistic variables and the preferred antecedent. More details are offered below.

Study 1: Oral production task (Narratives)

Two discourse contexts were taken into account: TC and TS. Multinomial logistic regressions were conducted in each context to analyse the association of Category of subjects (NS, LS, OSP) considered as a dependent variable with sociolinguistic variables (Group of speakers, Age, Proficiency) considered as independent variables (predictors). The relative risk ratio (RR) was used to establish positive association (RR>1), null association (RR=1) or negative association (RR<1) between dependent and independent variables. The measure RR indicated the likelihood that a subject occurred in a particular context as compared to a baseline.

Four multinomial logistic regression models were analysed for each of the two contexts (TC, TS). Table 5.18 summarises the four models. The first regression model used Group of speakers and Age as independent variables considering Spanish and Greek monolinguals. The second model used Group of speakers and Age as independent variables considering only the Greek-speaking groups. The third model used Group of speakers and Age as independent variables considering only the bilingual groups. The fourth model used Group of speakers, Age and Proficiency as independent variables considering only the groups of HS and L2ers.

Table 5.18. Multinomial logistic regression models analysed in TC and TS in narratives

Regression Model	Dependent variables	Independent variables
First model: Spanish and Greek monolinguals	Category of subjects (NS, LS, OSP)	• Group of speakers (Spanish, Greek) • Age at testing
Second model: Greek-speaking groups	Category of subjects (NS, LS, OSP)	• Group of speakers (Greek, IMM, HS, L2) • Age at testing
Third model: Bilingual groups	Category of subjects (NS, LS, OSP)	• Group of speakers (IMM, HS, L2) • Age at testing
Fourth model: HS and L2ers	Category of subjects (NS, LS, OSP)	• Group of speakers (HS, L2) • Age at testing • Proficiency

Study 2: Interpretation task (Anaphora resolution)

In AR, there were two contexts of TC (DDN, DIN) and two contexts of TS (DDO, DIO). The participants' responses indicated the preferred coreference of the embedded subject (NS, OSP) with a matrix antecedent (subject, object). Binomial logistic regressions were used to analyse the association of Antecedent preference (AP) considered as a dependent variable with sociolinguistic variables (Group of speakers, Age, Proficiency) considered as independent variables (predictors). Odds ratios (OR) established positive association (OR>1), null association (OR=1) or negative association (OR<1) between dependent and independent variables. The measure OR indicated the likelihood that an antecedent was selected as compared to a baseline.

Four binomial logistic regression models analysed the AR data for each condition (Table 5.19). The first model used Group of speakers and Age as independent variables considering Spanish and Greek monolinguals. The second model used Group of speakers and Age as independent variables considering the Greek-speaking groups. The third model used Group of speakers and Age as independent variables considering the bilingual groups. The fourth model used Group of speakers, Age and Proficiency as independent variables considering the groups of HS and L2ers.

Table 5.19. Binomial logistic regression models analysed in TC and TS in AR

Regression Model	Dependent variables	Independent variables
First model: Spanish and Greek monolinguals	Antecedent preference (subject, object)	· Group of speakers (Spanish, Greek) · Age at testing
Second model: Greek-speaking groups	Antecedent preference (subject, object)	· Group of speakers (Greek, IMM, HS, L2) · Age at testing
Third model: Bilingual groups	Antecedent preference (subject, object)	· Group of speakers (IMM, HS, L2) · Age at testing
Fourth model: HS and L2ers	Antecedent preference (subject, object)	· Group of speakers (HS, L2) · Age at testing · Proficiency

In two cases, the association between Age and AR responses was further explored through ANOVA only to verify the statistical significance of the association. This analysis was performed using the AR preferences of the participants as independent variables and Age as the dependent variable, following a reversed strategy from that of the main regression model established in the analysis (see §7.3.1.1, §7.3.1.2).

The results of the analyses of the production and interpretation data are presented and discussed in Chapter 6 and Chapter 7 respectively.

6. Study 1: Narratives

The present chapter concerns the oral production task. It begins by formulating the research questions and predictions in Section 6.1. Subsequently, it reports on the results drawn from the production data. Section 6.2 starts with an overview of the group data including descriptive and inferential statistics to provide the primary picture of the narratives corpus. Section 6.3 concentrates on the discourse-pragmatic use of subjects, which is the main focus of the study. A qualitative analysis of the data also follows each section of quantitative analysis in a case-by-case empirical approach. The role of the ambiguous verb morphology in the distribution of subjects in the Spanish production data is also examined. Section 6.4 reports on the regression analyses of the monolingual and bilingual production data. Section 6.5 offers a summary and Section 6.6 discusses the findings.

6.1 Research questions and predictions

Based on previous research evidence and theoretical accounts, the research questions and the predictions guiding Study 1 are formulated in the following sections.

6.1.1 First research question: Greek and Spanish monolinguals

Is Greek different from Spanish in the distribution of third-person null and overt subjects? If so, in which contexts and what causes divergence?

What is the effect of age at testing on the production of subjects by monolinguals?

Rationale for the predictions

Spanish and Greek are similar but not identical NS languages; therefore they should not be equivalent in the distribution of third-person subjects. Differences are expected in the relative frequency and/or the scope of OSP while the two languages should be similar in the scope of NS.

There seem to be differences in the scope of third-person OSP in the two languages suggesting the possibility of a more widespread use of OSP in Spanish than in Greek. This hypothesis is based on the following rationale:

(a) In Greek, the third-person OSP is deictically marked because it is the demonstrative which can assume anaphoric function, whereas in Spanish the personal pronoun has not such status. The deictic nature of the Greek pronoun renders its use less frequent as it is more discoursally marked than the Spanish one.

(b) Spanish syncretism in the inflectional marking of subject person renders verb morphology ambiguous in certain paradigms, which could lead to use of OSP for person disambiguation purposes regardless of context. In Greek, there is no verb inflectional ambiguity which would trigger OSP use to the same extent (see §2.3.7).

(c) Previous research on Greek and Spanish AR indicates that in neutral (non-focused) contexts the Spanish OSP establishes coreference to subjects more often than the Greek OSP. This suggests that in Spanish the OSP is more easily used in TC, hence it is more variable than in Greek; consequently, its scope seems to be wider than in Greek.

Predictions

Monolingual Greek and Spanish performance is expected to be in accordance with accessibility models, which may apply with flexibility because of stylistic choices and presence of contextual, grammatical and semantic cues. Specifically, in oral production, monolingual speakers are expected to produce contextualised sequences of sentences in narratives mainly encoding TC with NS and TS with LS or OSP. The frequency of NS is predicted to be significantly higher than that of LS and OSP. Due to general pragmatic and economy principles, monolingual groups are expected to use non-ambiguous NS in TS contexts, but less so than in contexts of TC.

Differences between the two languages are expected to arise in the use of OSP. The use of NS is expected to be relatively similar, with Greek speakers possibly using more NS in TS contexts. Greek monolinguals are not expected to use OSP in TC except in cases of focus. Spanish monolinguals are expected to use non-focused OSP in TC to a small extent since tolerance of OSP redundancy has been attested in

previous research. Although distribution of NS and overt subjects would be generally similar in Greek and Spanish, their scope and relative frequency should be to some extent different. In particular, the Greek OSP is expected to be relatively rare compared to its Spanish counterpart manifesting a narrower scope.

With regard to the role of age, in line with Hendriks et al. (2014), older monolingual speakers may be prone to potentially ambiguous referential choices due to age-related decline in cognitive capacity. This entails less sensitivity in tracking the discourse prominence of reference compared to younger adults. Both groups of monolinguals thus could show some instances of under-expliciteness produced by older speakers.

6.1.2 Second research question: Greek-Spanish bilinguals

Is Greek in contact with Spanish different from monolingual Greek in the distribution of third-person null and overt subjects? If so, in which contexts and what causes divergence?

What is the effect of age at testing on the production of subjects by bilinguals?

Rationale for the predictions

Greek in contact with Spanish was expected to be different from monolingual Greek in the distribution of third-person subjects. Differences were expected in the relative frequency and/or the scope of OSP but also, to a lesser degree, in the relative frequency and/or the scope of NS pronouns.

It is assumed that there is some residual activation of the other language in bilinguals at all times (e.g. Grosjean 2008) and crosslinguistic differences are predicted to exist in the two languages. Additionally, subject distribution in both languages is governed by interface conditions. In bilingual Greek, the interface-conditioned use of OSP was expected to be overgeneralised to discourse pragmatically inappropriate contexts due to language contact with Spanish under reduced input conditions. Moreover, overgeneralisation of NS was also possible, as evidenced in previous research (§3.7).

Overextension of the scope of both OSP and NS in bilingual Greek is predicted to be found on the grounds of the following rationale:

(a)   The distribution of pronominal subjects is constrained by discourse-pragmatic factors. The syntax-discourse/pragmatics interface is a vulnerable area in language contact situations because it requires simultaneous integration of syntactic and discourse/contextual information. This causes processing cost in bilinguals, who appear to treat the stronger element, i.e. the OSP, as the default resulting in overuse and/or over-acceptance of OSP regardless of language combination (Sorace et al. 2009; Sorace 2011, 2012; Tsimpli 2011).

(b)   OSP in NS languages, such as Greek and Spanish, are said to be pre-eminently problematic in language contact conditions due to their location at the syntax-discourse/pragmatics interface (e.g. Sorace 2011). In addition, NS can also be referentially complex involving the syntax-discourse/pragmatics interface, which may result in misuse and/or misinterpretation of NS (Clements & Domínguez 2016).

(c)   Despite the fact that Greek and Spanish are consistent NS languages, there seem to be crosslinguistic differences in the distribution of subject pronouns, which may trigger crosslinguistic influence. The scope of OSP in Spanish is wider than in Greek by hypothesis. In addition, Spanish is the dominant language in the context of the current study. Since the direction of crosslinguistic influence may also depend on language dominance (e.g. Argyri & Sorace 2007), crosslinguistic effects may surface in some overgeneralisation of OSP in Greek due to influence from Spanish. In other words, if the answer to the first research question will be positive, i.e. if OSP in Spanish were not constrained by exactly the same discourse/pragmatic restrictions as in Greek, crosslinguistic influence could be a possible source of variation in bilingual Greek. On the other hand, no systematic differences were expected on the scope of NS between the two languages; consequently, it cannot be assumed that there is chance of crosslinguistic influence at the level of NS. If differences appear in the scope of NS in the bilingual performance, as in other studies (§3.7), these could be attributed to interface conditions/processing factors corroborating the referential complexity of NS (Clements & Domínguez 2016).

(d)   In sum, overgeneralisation of both OSP and NS was predicted in bilinguals due to more taxed processing resources stemming from interface conditions. This

could be reinforced by crosslinguistic influence from Spanish to Greek in the case of OSP. Therefore, bilingual Greek was expected to differ from monolingual Greek.

(e)   Differences initiate in first-generation immigrant speakers (immigrants) with stronger effects in HS and L2ers. In line with Kaltsa et al. (2015: 269), 'a change from what was once a monolingual grammar as is the case of L1 attrition should show less divergence than a bilingual grammar that has followed a different developmental path and a different endstate than that of a monolingual grammar'.

Predictions

In oral production, Greek-Spanish bilinguals were expected to produce connected speech in narratives and to generally use referring expressions in the same way as monolinguals, i.e. using NS mostly in TC and overt subjects (LS, OSP) mostly in TS.

It has been reported that HS tend to avoid embedded structures (Polinsky & Kagan 2007; Polinsky 2008; Benmamoun et al. 2013). This may be also the case for lower proficiency L2ers due to fluency constraints. In the HS' production, the number of embedded clauses could be lower than in monolingual production, possibly depending on language proficiency. Low use of embeddings may affect the relative frequency of NS and overt subjects. Embedded clauses favour the use of NS, whereas matrix clauses do not favour either form (Silva-Corvalán 1994; Margaza & Bel 2006; Prada Pérez 2009). Fewer embedded clauses imply higher numbers of overt subjects and lower numbers of NS compared to monolingual/native discourse. Additionally, since it has been consistently found that L2ers are more explicit (over-explicit) in reference maintenance than monolinguals (Hendriks 2003), overproduction of overt subjects was anticipated in the L2 performance compared to the other groups.

Overall, in Greek in contact with Spanish, overuse of overt subjects was expected in contexts of TC by HS and L2ers compared to monolinguals. Moreover, overuse of OSP was expected to some extent in the bilingual performance in both TC and TS due to interface conditions, as in previous studies in pairs of two NS languages (§3.4), but possibly also due to crosslinguistic influence. Overuse of NS in contexts of TS might also obtain, with ambiguous NS constructions being possible (§3.7, §3.8).

In line with Hendriks et al. (2014) and Kaltsa et al. (2015), age may play a significant role in the linguistic behaviour of bilinguals with respect to reference. In particular, the older bilingual speakers may be more susceptible to producing ambiguous reference due to age effects on general cognitive skills, which could be enhanced by bilingualism. This suggests that in the groups of bilinguals the older speakers would produce instances of ambiguity to a greater extent than monolinguals and younger bilinguals. The potential age effect in the participants' production of subjects is examined through logistic regression analyses.

6.2 Overview of the results

This section offers an overview of the production data drawn from oral narratives. The basic unit of analysis was the clause, i.e. each utterance containing a verb and a referential subject. A total of 232 (=116*2) narratives based on the story telling tasks were analysed after being transcribed, coded and annotated. All narratives were in Greek except those produced by the Spanish monolinguals (see details in Chapter 5).

6.2.1 Length of narratives

Table 6.1 shows the length of narratives produced by the participants considering both stories as per number of clauses uttered by each speaker in each group. In the monolinguals, the shortest narrative reached a length of 26 clauses in both languages while the maximum length was 61 clauses for the Spanish and 80 clauses for the Greek monolinguals. For the immigrants, the range was 25-65 clauses per speaker while the HS and the L2ers produced shorter narratives with ranges of 17-64 and 22-49 clauses respectively. The mean number of clauses per group shows that the speakers for whom the language of narratives (Greek) was their weak language (HS, L2) were more reserved in speech than the monolingually raised groups (monolinguals, immigrants).

Table 6.1. Length of narratives (N of clauses) per group of speakers

Group	N of speakers	N of clauses per speaker			
		Mean	**SD**	**Min**	**Max**
Spanish	20	44.5	9.4	26	61
Greek	20	49.1	15.5	26	80
IMM	35	45.0	10.8	25	65
Younger IMM	23	45.0	11.4	25	65
Older IMM	12	45.0	10.1	34	65
HS	21	32.3	10.5	17	64
L2	20	35.8	7.8	22	49

As depicted in Table 6.1, the mean and SD length of narratives differ between groups, with HS and L2 producing considerably shorter narratives than the others. This is also illustrated in Figure 6.1 with one outlier in the HS group. A one-way ANOVA showed that the monolingually raised groups speaking in their L1 (considered as one group) gave significantly longer narratives as per number of clauses than HS and L2 (considered as another group) speaking in their weak language [$F_{(1,4837)}=1202.75$, $p<0.001$].
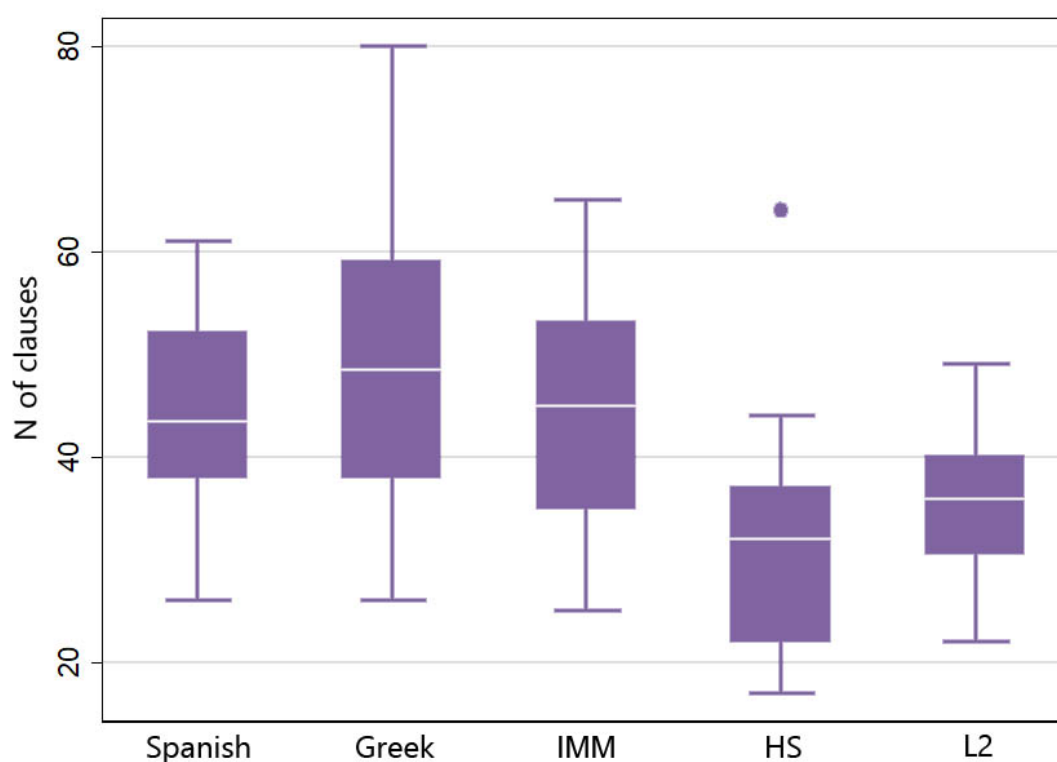
Figure 6.1. Box plot of number of clauses per participant per group of speakers

A one-way ANOVA showed that proficiency level (basic, intermediate, advanced, near native) was not statistically associated with the length of narratives in the HS and L2 groups [F (3,37)=0.45, p=0.719].

Table 6.2 shows the N of clauses produced by each group of speakers indicating the N of null and overt subjects as well as that of subject-headed relative clauses (RC). This type of RC (N=322, 6.65%) were separated from the corpus (as in Montrul & Rodríguez Louro 2006; Shin 2012) to be examined in future work. Despite differences in narratives' length, all groups were generally similar in subject distribution as indicated by the percentages, which are given in slanted numbers in this and all the subsequent tables.

Table 6.2. Total N and percentage of clauses per group of speakers in narratives

	Spanish	Greek	IMM	HS	L2	Total
Null	503	606	944	387	365	2,807
	56.52	*61.71*	*59.98*	*57.08*	*51.05*	*58.01*
Overt	314	330	515	246	307	1712
	35.28	*33.60*	*32.72*	*36.28*	*42.94*	*35.38*
RC	73	46	115	45	43	322
	8.20	*4.68*	*7.31*	*6.64*	*6.01*	*6.65*
Total	890	982	1,574	678	715	4,839
	100	*100*	*100*	*100*	*100*	*100*

Apart from group results, individual results are additionally considered whenever necessary to detect biased results potentially produced by individual performance.

6.2.2 Clauses

Matrix clauses were used more frequently than embedded clauses (Figure 6.2, Table 6.3). Spanish monolinguals produced matrix and embedded clauses to a similar extent. The Greek-speaking groups used significantly more matrix than embedded clauses. The HS did not produce significantly fewer embedded clauses than Greek monolinguals or immigrants (cf. Polinsky & Kagan 2007; Polinsky 2008). The L2ers produced significantly fewer embedded clauses than the immigrants [Pearson χ^2 (1, N=2,289)=8.302, p=0.004], but they did not differ from Greek monolinguals or HS.
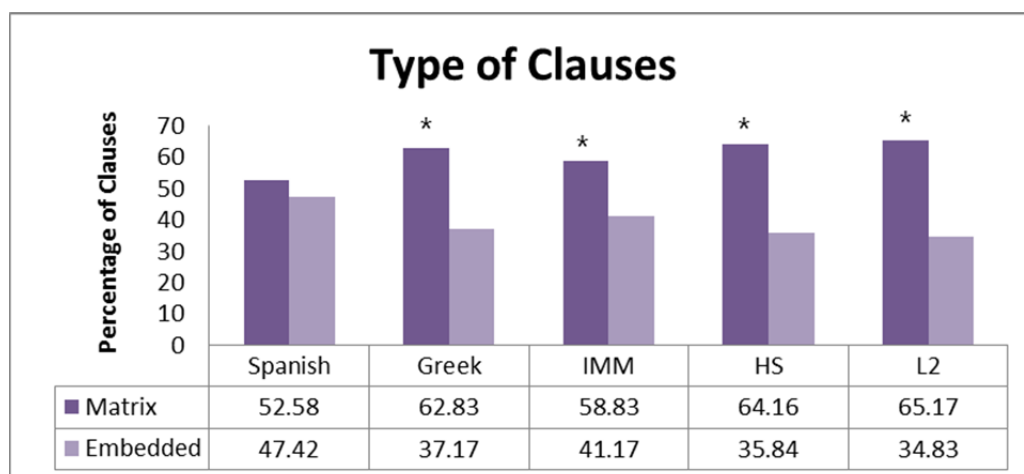
Figure 6.2. Type of clauses in narratives

Table 6.3. Type of clauses in narratives

	Spanish	Greek	IMM	HS	L2	Total
Matrix	468	617	926	434	466	2,911
	52.58	*62.83*	*58.83*	*64.01*	*65.17*	*60.18*
Embedded	422	365	648	244	249	1,928
	47.42	*37.17*	*41.17*	*35.99*	*34.83*	*39.84*
Total	890	982	1,574	678	715	4,839
	100	*100*	*100*	*100*	*100*	*100*

The use of matrix and embedded clauses was associated with proficiency level in HS and L2ers. A binomial logistic regression (N of observations: 1393) showed a significant (p=0.009) and positive association (OR=1.156) between Type of clause and Proficiency, indicating that being more proficient by one level increases the chance of using embedded clauses by 15.6% (Table 6.4).

Table 6.4. Binomial logistic regression: type of clause and proficiency

Type of Clause	OR	SE	z	p	OR CI	
					LB	UB
Proficiency	1.156	0.064	2.60	0.009	1.036	1.289

(OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)
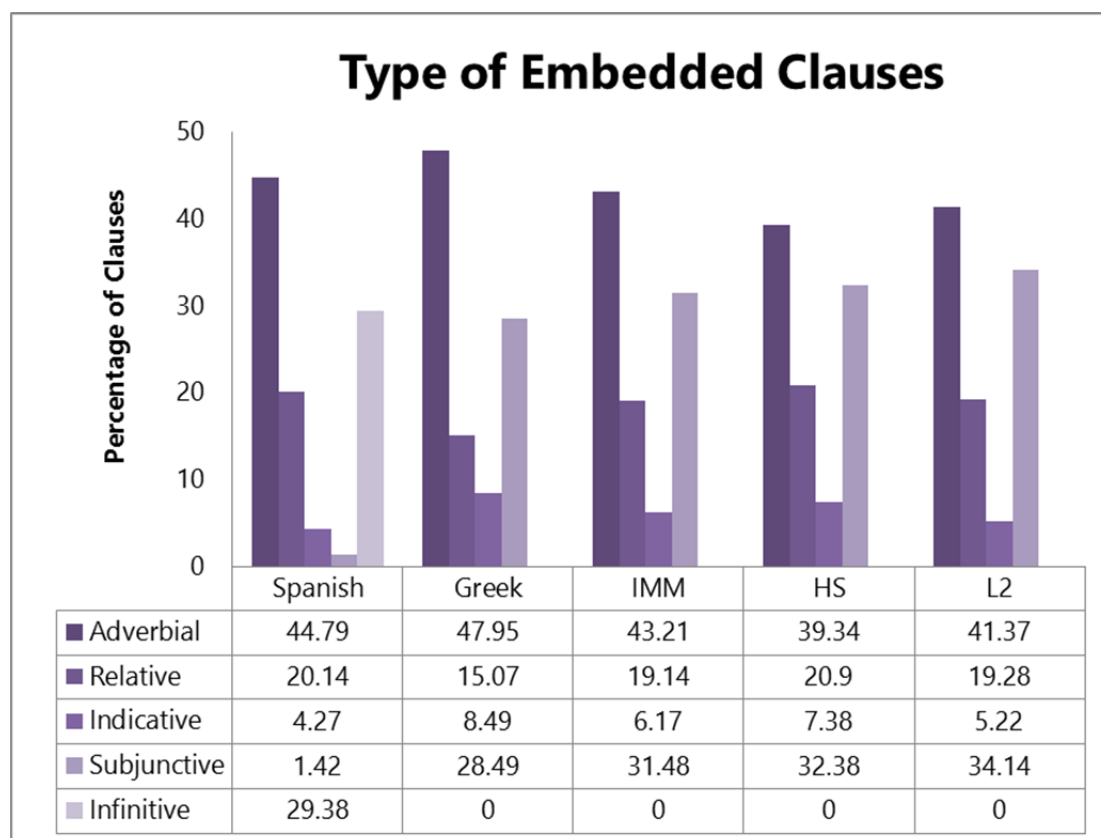
Figure 6.3. Type of embedded clauses in narratives

Table 6.5. Type of embedded clauses in narratives

	Spanish	Greek	IMM	HS	L2	Total
Adverbial	189	175	280	96	103	843
	44.79	*47.95*	*43.21*	*39.34*	*41.37*	*43.72*
Relative	85	55	124	51	48	363
	20.14	*15.07*	*19.14*	*20.9*	*19.28*	*18.83*
Indicative Compl.	18	31	40	18	13	120
	4.27	*8.49*	*6.17*	*7.38*	*5.22*	*6.22*
Subjunctive Compl.	6	104	204	79	85	478
	1.42	*28.49*	*31.48*	*32.38*	*34.14*	*24.79*
Infinitive Compl.	124	0	0	0	0	124
	29.38	*0*	*0*	*0*	*0*	*6.43*
Total	422	365	648	244	249	1,928
	100	*100*	*100*	*100*	*100*	*100*

Figure 6.3 and Table 6.5 show the type of embedded clauses used in narratives and their rates. Complement clauses were those refered to as indicative, subjunctive and infinitive in the Figure 6.3. All groups of speakers produced mostly adverbial clauses. The greatest difference was the high percentage of subjunctive complement clauses

in the Greek-speaking groups compared to Spanish (28.5%, 31.5%, 32.4%, 34.1% vs 1.4% in Spanish). The proportion of subjunctive complement clauses in Spanish was low due to the extensive use of the infinitive (29.4%), which was absent from the Greek data (see Goodluck et al. 2001; Parodi & Tsimpli 2005; Spyropoulos 2007). This difference did not affect subject distribution in monolinguals since NS rates in infinitives and subjunctive complement clauses were by far predominant while use of overt subjects was scarce.

6.2.3 Subjects

This section offers an overview of the use of subjects in all groups with regard to subjects' category (LS, OSP, NS, relative pronoun). Pearson Chi-square tests were used to detect significant differences between the groups and two-sample proportion tests were used to explore significant differences within groups.

6.2.3.1 Subjects in all clauses

Figure 6.4 and Table 6.6 show that more than 50% of subjects were omitted considering all clauses (matrix, embedded) and all contexts (focus, TC, TS). Greek monolinguals manifested the greatest rate of NS (61.71%), showing a statistically significant difference from that of Spanish (56.52%) [Pearson χ^2 (1, N=1,872)=5.21, p=0.022] and the L2 group (51.05%) [Pearson χ^2 (1, N=1,697)=19.21, p<0.001]. LS were used at a similar rate by all groups (32%-35.55%), excepting the L2 group, which manifested a greater percentage of LS (41.54%), being statistically different from monolingual Greek [Pearson χ^2 (1, N=1,697)=13.01, p<0.001], immigrants [Pearson χ^2 (1, N=2,294)=18.95, p<0.001] and HS [Pearson χ^2 (1, N=1,393)=5.27, p=0.022].

The frequency of OSP was very low (0.44%-2.7%) compared to NS and LS. Spanish monolinguals used significantly more OSP (2.7%) than Greek monolinguals (0.61%) [Pearson χ^2 (1, N=1,872)=12.87, p<0.001], immigrants (0.44%) [Pearson χ^2 (1, N=2,464)=23.20, p<0.001] and HS (0.7%) [Pearson χ^2 (1, N=1,568)=8.13, p=0.004]. The L2 group was not different from either monolingual group in OSP rates, but was

statistically different from immigrants [Pearson χ^2 (1, N=2,289)=6.06, p=0.014]. No other statistically significant differences were detected in the Greek-speaking groups.
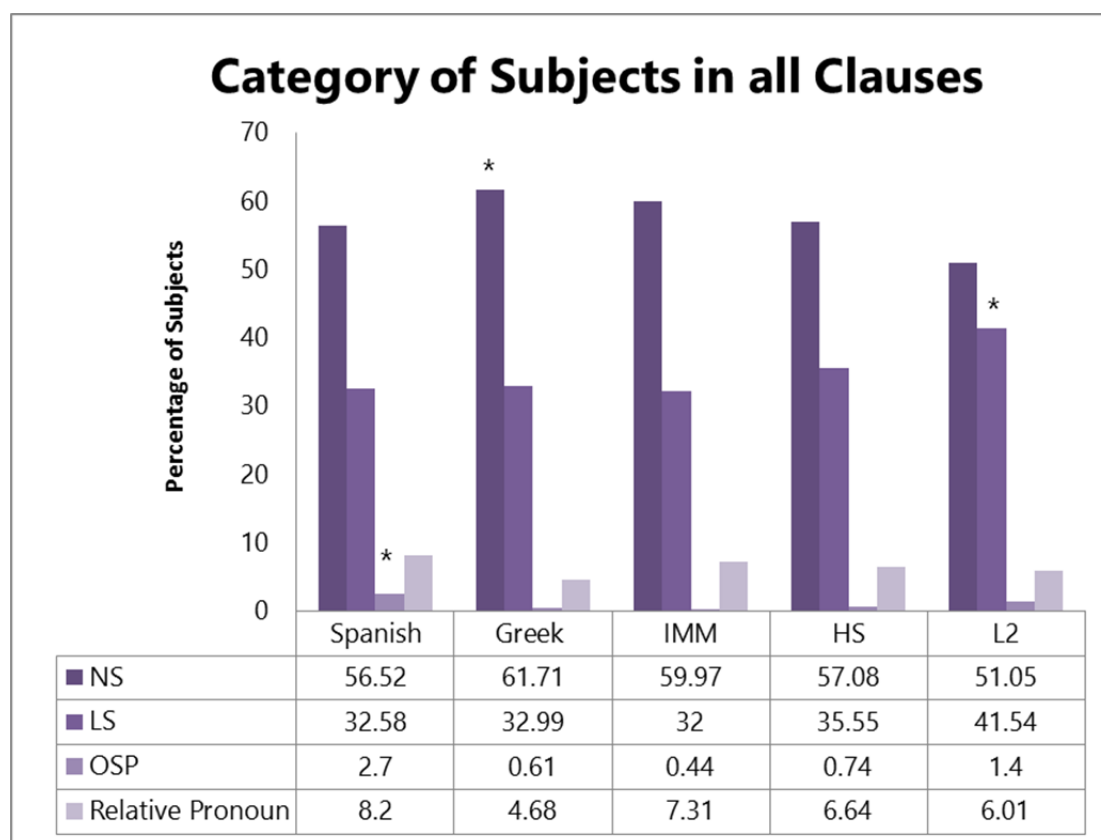
Figure 6.4. Category of subjects in all clauses in narratives

Table 6.6. Category of subjects in all clauses in narratives

	Spanish	Greek	IMM	HS	L2	Total
Null Pronoun	503	606	944	387	365	2,805
	56.52	*61.71*	*59.97*	*57.08*	*51.05*	*57.97*
Lexical Subject	290	324	508	241	297	1,660
	32.58	*32.99*	*32*	*35.55*	*41.54*	*34*
Overt Pronoun	24	6	7	5	10	52
	2.7	*0.61*	*0.44*	*0.74*	*1.4*	*1.07*
Relative Pronoun	73	46	115	45	43	322
	8.2	*4.68*	*7.31*	*6.64*	*6.01*	*6.65*
Total	890	982	1,574	678	715	4,839
	100	*100*	*100*	*100*	*100*	*100*

In Spanish, 5 out of 24 OSP were instances of the demonstrative *este* used in focus (N=2) and TS (N=3). In monolingual Greek, 2 out of the 6 occurrences of OSP were instances of the demonstrative *ekinos*. The latter pronoun was also used twice in the immigrants and once in the HS group (see §6.3.1.2 and §6.3.2.2 for details).

6.2.3.2 Subjects in matrix clauses

Considering only the matrix clauses (Figure 6.5, Table 6.7), Greek monolinguals, immigrants, and HS used NS more than 50% of the time. Spanish and L2ers used NS at lower rates (45.09%, 43.13% respectively). Two-sample tests of proportions showed that monolingual Greeks used NS significantly more than LS (z=2.227, p=0.025), while in Spanish the difference between the percentages of NS and LS was not statistically significant. There was a statistically significant difference in the frequences of NS between Spanish and Greek [Pearson χ^2 (1, N=1,085)=9.35, p=0.002] as well as between their frequences of LS [Pearson χ^2 (1, N=1,085)=4.06, p=0.044].

Category of Subjects in Matrix Clauses

	Spanish	Greek	IMM	HS	L2
■ NS	45.09	54.46	52.38	51.72	43.13
■ LS	51.07	44.89	47.08	47.36	55.36
■ OSP	3.85	0.65	0.54	0.92	1.5
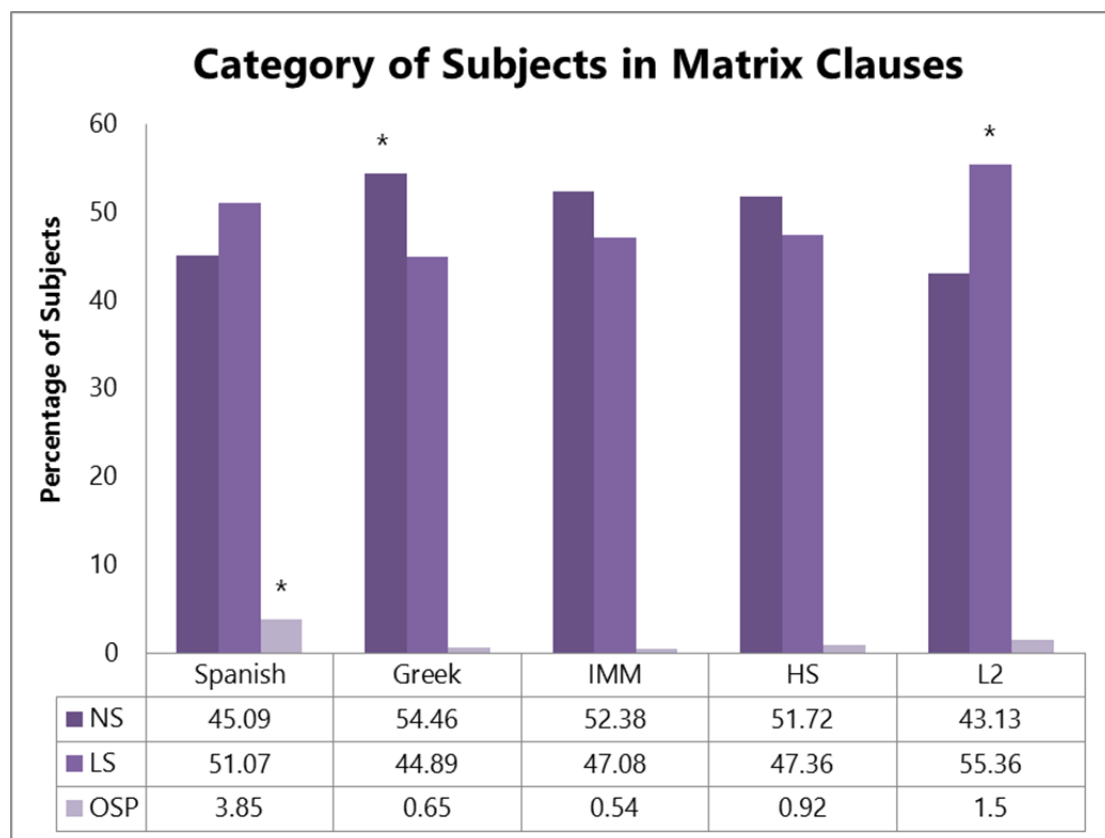
Figure 6.5. Category of subjects in matrix clauses in narratives

The L2 group produced LS significantly more than NS (z=-2.599, p=0.009) performing similarly to Spanish, but differently from monolingual Greek in LS use [Pearson χ^2 (1, N=1,083)=11.64, p=0.001] and in NS use [Pearson χ^2 (1, N=1,083)=13.61, p<0.001]. Immigrants and HS used NS and LS to similar rates, performing similarly to Greek monolinguals.

The frequency of OSP, overall much lower relative to NS and LS, was higher in Spanish compared to Greek (3.85% vs 0.65%), this difference reaching statistical significance [Pearson χ^2 (1, N=1,085)=13.70, p<0.001]. The rate of OSP was also higher in Spanish relative to the L2 group (3.85% vs 1.5%) [Pearson χ^2 (1, N=934)=4.92, p=0.026] (this was not the case when considering matrix and embedded clauses together). The four Greek-speaking groups performed similarly in OSP production in matrix clauses and significantly differently from the Spanish group.

Table 6.7. Category of subjects in matrix clauses in narratives

	Spanish	Greek	IMM	HS	L2	Total
Null Pronoun	211	336	485	224	201	1,457
	45.09	*54.46*	*52.38*	*51.61*	*43.13*	*50.05*
Lexical Subject	239	277	436	206	258	1,416
	51.07	*44.89*	*47.08*	*47.47*	*55.36*	*48.64*
Overt Pronoun	18	4	5	4	7	38
	3.85	*0.65*	*0.54*	*0.92*	*1.5*	*1.31*
Total	468	617	926	434	466	2,911
	100	*100*	*100*	*100*	*100*	*100*

6.2.3.3 Subjects in embedded clauses

Considering only the embedded clauses, NS were much more frequent than LS and OSP in all groups in more than 65% (Figure 6.6, Table 6.8). The relative frequency of NS in Greek monolinguals was found to be significantly higher compared to the L2 group (73.97% vs 65.86%) [Pearson χ^2 (1, N=614)=4.69, p=0.030] and it did not significantly differ from the Spanish group. The LS rate in the L2 group was the highest of all (15.66%) without significantly differing from the other groups' rates. The use of OSP was not found to be significantly different between Spanish and Greek monolinguals. In all Greek-speaking groups (monolinguals, IMM, HS, L2), the relative frequency of OSP was roughly the same. Subject-headed relative clauses were produced at similar rates in all groups (17.3%-18.44%), except for Greek monolinguals presenting a relatively low percentage (12.6%).

Category of Subjects in Embedded Clauses

	Spanish	Greek	IMM	HS	L2
■ NS	69.19	73.97	70.83	66.8	65.86
■ LS	12.09	12.88	11.11	14.34	15.66
■ OSP	1.42	0.55	0.31	0.41	1.2
■ Relative Pronoun	17.3	12.6	17.75	18.44	17.27
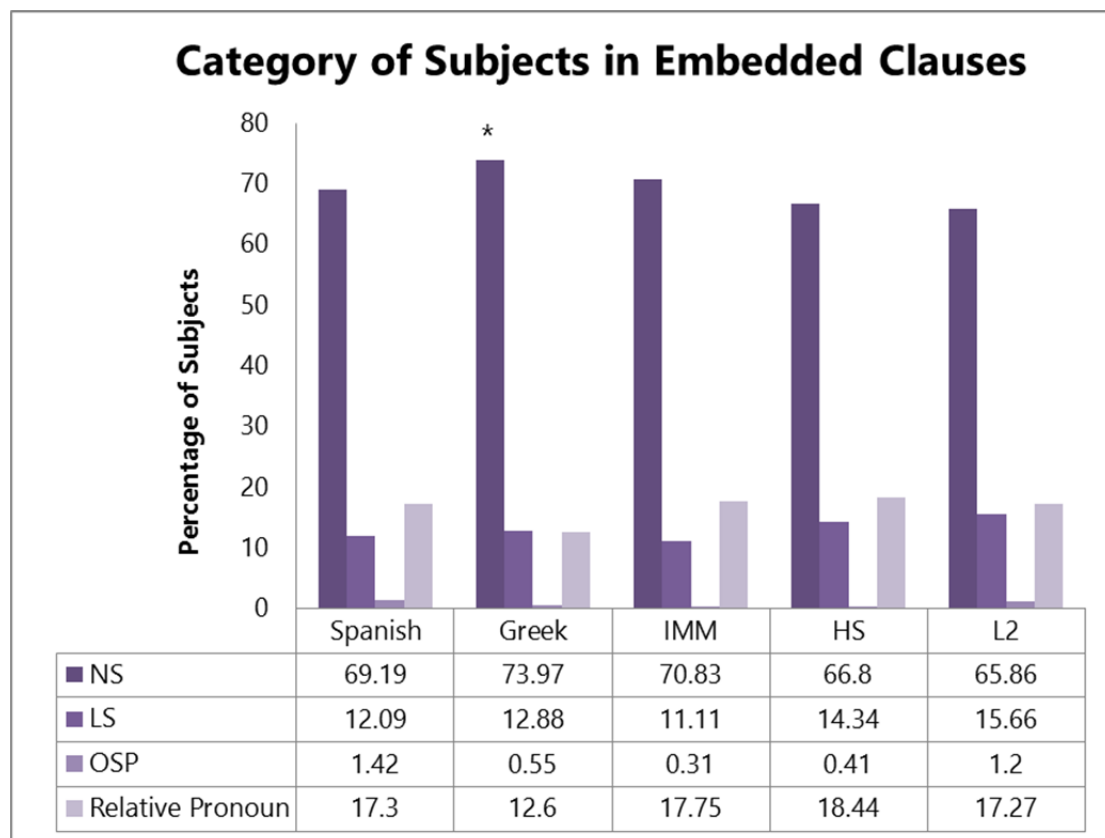
Figure 6.6. Category of subjects in embedded clauses in narratives

Table 6.8. Category of subjects in embedded clauses in narratives

	Spanish	Greek	IMM	HS	L2	Total
Null Pronoun	292	270	459	163	164	1,348
	69.19	*73.97*	*70.83*	*66.80*	*65.86*	*69.92*
Lexical Subject	51	47	72	35	39	244
	12.09	*12.88*	*11.11*	*14.34*	*15.66*	*12.66*
Overt Pronoun	6	2	2	1	3	14
	1.42	*0.55*	*0.31*	*0.41*	*1.2*	*0.73*
Relative Pronoun	73	46	115	45	43	322
	17.3	*12.6*	*17.75*	*18.44*	*17.27*	*16.70*
Total	422	365	648	244	249	1,928
	100	*100*	*100*	*100*	*100*	*100*

6.3 Discourse-pragmatic use of subjects

The present section concerns the discourse-pragmatic use of subjects in narratives, which reveals patterns of subject use in particular contexts. When a referent was introduced into the discourse for the first time in subject position, the subject referent was tagged as a being focused (new information). New referents are always realised as nouns (e.g. Paredes Silva 1993; Hendriks et al. 2014). If the new referent introduced in existential constructions and presentational sentences was one of the story participants (*horse*, *cow*, *bird*, *little birds*, *cat*, *dog*), then this was considered presentational focus. If the referent was not focused, it was considered to be topic (old information) in contexts of TC, i.e. subject referent maintenance (coreference); or TS, i.e. change of subject referent (non-coreference).
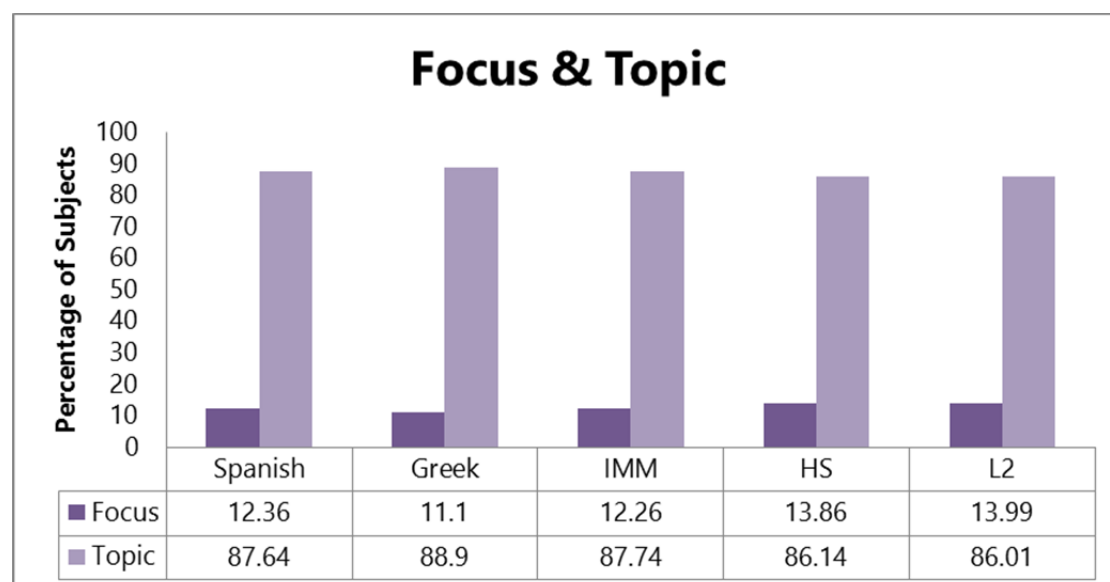
Figure 6.7. Subjects in focus and topic contexts in narratives

Table 6.9. Subjects in focus and topic contexts in narratives

	Spanish	Greek	IMM	HS	L2	Total
Focus	110	109	193	94	100	606
	12.36	*11.1*	*12.26*	*13.86*	*13.99*	*12.52*
Topic	780	873	1381	584	615	4,233
	87.64	*88.9*	*87.74*	*86.14*	*86.01*	*87.48*
Total	890	982	1574	678	715	4,839
	100	*100*	*100*	*100*	*100*	*100*

All groups produced comparable rates of subjects in focus and topic contexts (Figure 6.7, Table 6.9). Subject distribution was scrutinised in TC and TS discourse contexts. Clauses introducing focused subjects, i.e. 606 clauses (12.52%), were excluded keeping 4,233 topic clauses (87.48%) of the total. All *who*-RC (N=322, 7.6%) were also excluded. Therefore, 3,911 (92.39%) out of 4,233 were included in the main study. Table 6.10 lists the basic features of the subjects that were considered in the analyses.

Table 6.10. Features of subjects considered in the analyses

Features of subjects		%		
Number	93.4%	singular	6.6%	plural
Animacy	99%	animate	1%	inanimate
Definiteness (LS)	99%	definite	1%	indefinite
Position (LS, OSP)	73.3%	preverbal	26.7%	postverbal

Given the task storylines, the majority of subjects used in production were in third-person singular (93.4%) with plural emerging to a much lesser extent (6.6%). Both animate and inanimate subjects were analysed, with the vast majority thereof being animate (99%) and few inanimate ones (1%) (e.g. *fence*, *leg*, *nest*). Virtually all LS were definite (99%) as they appeared in topic contexts. Finally, 73.3% of overt subjects were preverbal and 26.7% postverbal (Spanish: 62.2% vs 37.8%; Greek: 52.1% vs 47.9%; immigrants: 53.6% vs 46.4%; HS: 54.1% vs 45.9%; L2ers: 62.2% vs 37.8% preverbal and postverbal subjects respectively).

Mismatch in subject-verb agreement was observed in HS (N=23, 4.38%) and L2ers (N=19, 3.35%). From the 42 instances of such mismatch, 40 (95.23%) were produced by basic and intermediate proficiency speakers. A binomial logistic regression (N of observations: 1371) showed a significant (p<0.001) and negative association (OR=0.152) between mismatch and proficiency. Being more proficient by one level decreases the chance of producing a mismatch by 6.6% (Table 6.11).

Table 6.11. Binomial logistic regression: mismatch and proficiency

Mismatch	OR	SE	z	p	OR CI	
					LB	UB
Proficiency	0.152	0.037	-7.59	<0.001	0.093	0.247

(OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval,
LB=Lower Bound, UB=Upper Bound)

6.3.1 Topic Continuity

6.3.1.1 Quantitative analysis

Starting with contexts involving referents' maintenance in subject position, the preferred way of expressing TC in all groups of speakers was by far using NS (86.36%-95.79%) (Figure 6.8, Table 6.12). All groups of speakers, including monolinguals, also used LS to express topic maintenance at a small rate.

In the use of LS, there was no statistical difference between the Spanish and Greek monolingual performance (5.91% vs 4.04%). The relative frequency of LS in TC was significantly lower in Greek monolinguals (4.04%) with respect to the L2 group (12.3%) [Pearson χ^2 (1, N=944)=22.76, p<0.001] and to HS (7.49%) [Pearson χ^2 (1, N=917)=5.098, p=0.024]. In addition, the relative frequency of LS in TC was significantly lower in immigrants (4.48%) compared to the L2 group [Pearson χ^2 (1, N=1,244)=25.10, p<0.001] and the HS [Pearson χ^2 (1, N=1,217)=4.44, p=0.035]. The L2 group used significantly more LS in TC than the HS [Pearson χ^2 (1, N=721)=4.62, p=0.031].

The percentage of NS was significantly higher in Greek monolinguals (95.79%) relative to HS (92.22%) [Pearson χ^2 (1, N=917)=5.23, p=0.022] and the L2 group (86.36%) [Pearson χ^2 (1, N=944)=27.43, p<0.001]. Similarly, in immigrants the percentage of NS (95.29%) was significantly higher compared to HS [Pearson χ^2 (1, N=1,217)=4.42, p=0.035] and the L2 group [Pearson χ^2 (1, N=1,244)=30.41, p<0.001].

The frequency of OSP in TC was very low in all groups with no major differences. It was actually significantly higher in the L2 group (1.34%) compared to monolingual Greek (0.18%) and immigrants (0.23%). However, the number of OSP was very low (N=5), produced by only two L2ers, one of whom produced 4 out of 5 OSP in TC. Thus, the statistical significance found in the results regarding OSP in TC in the L2 group is not conclusive (see also §6.3.1.2).
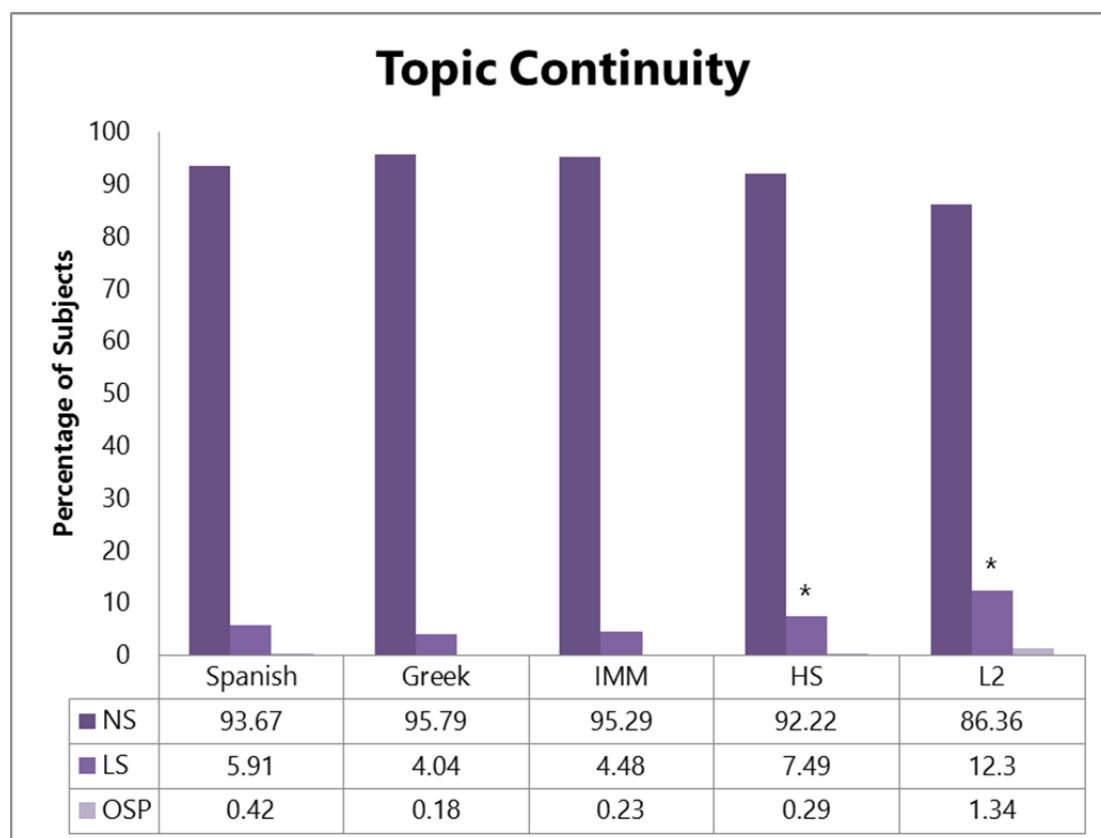
Figure 6.8. Use of subjects in TC contexts

Table 6.12. Use of subjects in TC contexts

	Spanish	Greek	IMM	HS	L2	Total
Null Pronoun	444	546	829	320	323	2,462
	93.67	*95.79*	*95.29*	*92.22*	*86.36*	*93.43*
Lexical Subject	28	23	39	26	46	162
	5.91	*4.04*	*4.48*	*7.49*	*12.3*	*6.15*
Overt Pronoun	2	1	2	1	5	11
	0.42	*0.18*	*0.23*	*0.29*	*1.34*	*0.42*
Total	474	570	870	347	374	2,635
	100	*100*	*100*	*100*	*100*	*100*

If we take into account only the matrix clauses (Figure 6.9, Table 6.13), the pattern of predominant use of NS is generally similar. Spanish and Greek monolinguals were statistically different regarding LS (11.59% vs 5.43%) [Pearson χ^2 (1, N=520)=6.51, p=0.011] and NS (87.44% vs 94.57%) [Pearson χ^2 (1, N=520)=8.34, p=0.004]. Additionally, Greek monolinguals used significantly fewer LS in TC in matrix clauses than HS (5.43% vs 10.34%) [Pearson χ^2 (1, N=516)=4.35, p=0.037] and L2ers (5.43% vs 17.45%) [Pearson χ^2 (1, N=525)=19.79, p<0.001]. Greek monolinguals also produced

161

significantly more NS in TC than HS (94.57% vs 89.16%) [Pearson χ^2 (1, N=516)=5.15, p=0.023] and L2ers (94.57% vs 81.13%) [Pearson χ^2 (1, N=525)=23.57, p<0.001]. The frequency of OSP was insignificant in all cases.
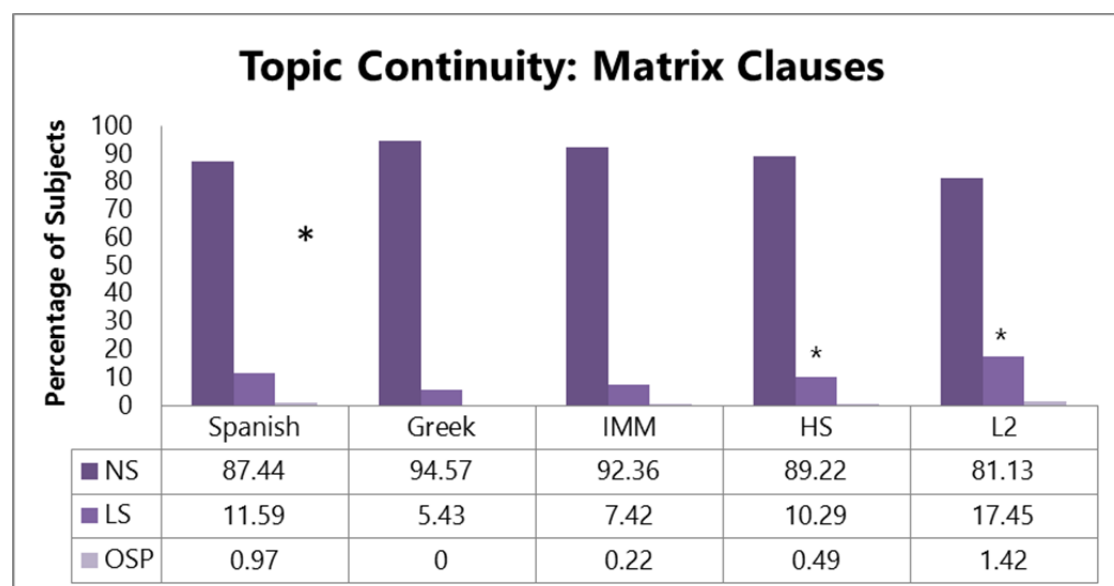
Figure 6.9. Use of subjects in TC contexts in matrix clauses

Table 6.13. Use of subjects in TC contexts in matrix clauses

	Spanish	Greek	IMM	HS	L2	Total
Null pronoun	181	296	411	181	172	1,241
	87.44	*94.57*	*92.36*	*89.16*	*81.13*	*89.93*
Lexical subject	24	17	33	21	37	132
	11.59	*5.43*	*7.42*	*10.34*	*17.45*	*9.57*
Overt pronoun	2	0	1	1	3	7
	0.97	*0*	*0.22*	*0.49*	*1.42*	*0.51*
Total	207	313	445	203	212	1,380
	100	*100*	*100*	*100*	*100*	*100*

If we take into account only the embedded clauses, NS were almost always used in TC (Figure 6.10, Table 6.14). NS rates were significantly lower in the L2 group compared to Greek monolinguals [Pearson χ^2 (1, N=419)=3.99, p=0.046] and immigrants [Pearson χ^2 (1, N=587)=10.43, p=0.001]. The use of LS in embedded TC clauses was very low and no significant differences were found between the groups. The frequency of OSP, as in the case of matrix clauses, was insignificant (see §6.3.1.1).
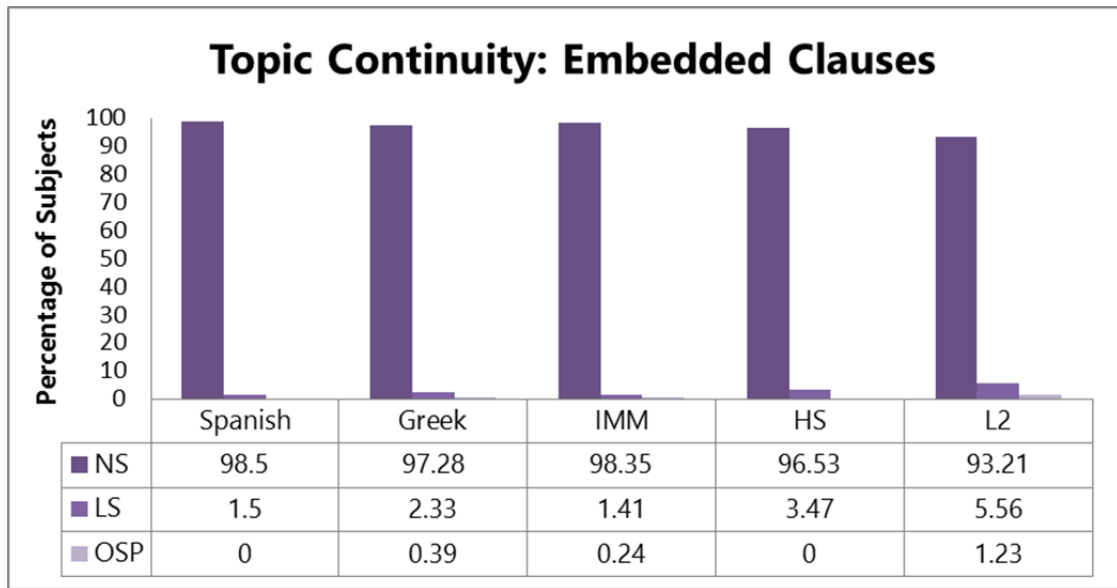
Figure 6.10. Use of subjects in TC contexts in embedded clauses

Table 6.14. Use of subjects in TC contexts in embedded clauses

	Spanish	Greek	IMM	HS	L2	Total
Null pronoun	263	250	418	139	151	1,221
	98.5	*97.28*	*98.35*	*96.53*	*93.21*	*97.29*
Lexical subject	4	6	6	5	9	30
	1.5	*2.33*	*1.41*	*3.47*	*5.56*	*2.39*
Overt pronoun	0	1	1	0	2	4
	0	*0.39*	*0.24*	*0*	*1.23*	*0.32*
Total	267	257	425	144	162	1,255
	100	*100*	*100*	*100*	*100*	*100*

Considering matrix and embedded clauses separately offers an insight into the ways in which extra-sentential and intra-sentential anaphora is realised with respect to subjects in semi-spontaneous production of oral narratives. In addition, a qualitative approach to the production data was essential, as shown in the next section.

6.3.1.2 Qualitative analysis

All groups of speakers used more NS than other categories of subjects in TC contexts and they also produced overt subjects in TC. Subjects' distribution in TC was scrutinised on the three categories of subjects used (NS, LS, OSP). A close look at the contexts of use sheds light into their (in)felicity.

Null subjects in topic continuity

All groups of speakers principally used NS in contexts of TC, such as in (38) and (39).

(38)   *Un gato observaba a los pajaritos. Al parecer ∅ quería hacerles daño, ∅ se los quería comer.*
       'A cat was observing the little birds. Seemingly, [he] wanted to harm them, [he] wanted to eat them.'

(S72, Spanish, age: 40)

(39)   *Enas skilos ftani ke ∅ piani tin ura tu- tis gatas ke tin ∅ petaksi ekso.*
       'A dog arrives and [he] catches the cat's tail and [he] throws her out.'

(S141, L2, age: 24, intermediate)

(Over)use of lexical subjects in topic continuity

All groups of speakers used LS in TC, which may be overly informative, hence infelicitous. This is because TC involves subject maintenance, thus overt subjects are not required. There were 162 cases of LS in TC (6.15%) overall (see §6.3.1.1). Greek and Spanish monolinguals performed similarly in the use of LS in TC. HS used significantly more LS in TC than Greek monolinguals and immigrants. The L2 group also used significantly more LS in TC than Greek monolinguals, immigrants and HS. Table 6.15 shows the N and percentage of LS in TC as well as the number and percentage of speakers who produced such instances and the maximum N of LS in TC per speaker.

164

Table 6.15. Use of lexical subjects in TC contexts in all groups

Group	LS in TC		Speakers using LS in TS		Maximum N of LS in TC per speaker
	N	%	N	%	N
Spanish	28/474	*5.91*	15/20	*75*	4
Greek	23/570	*4.04*	13/20	*65*	3
IMM	39/870	*4.48*	23/35	*65.7*	5
HS	26/347	*7.49*	13/21	*61.9*	4
L2	46/374	*12.3*	17/20	*85*	5

A qualitative analysis showed that not all cases of LS in TC were infelicitous. According to the particular context, two main categories could be distinguished: those that were (potentially) redundant and those that were not. The following subsections focus on (a) non-redundant and (b) redundant LS in TC.

(a) Non-redundant LS

Non-redundant LS in TC (N=125, 77.16%) were divided into the following types:

(i) LS following existential-presentational clauses

(ii) LS preceded by coreferential NS

(iii) LS forming equivalents

(iv) LS in focus

(v) LS in rephrasing

In these contexts, the LS were not infelicitous, as will be illustrated in examples. With regard to the frequency of this type of LS in TC per group in the production corpus, information can be seen in Table 6.16 at the end of this section.

(i) LS following existential-presentational clauses:

More than half of LS in TC contexts (87 out of 162) concerned LS in a TC clause following a clause with a focused subject, i.e. containing a new subject-character introduced in the narrative usually with an indefinite LS, found in the beginning of the narratives or episodes. The preceding subject was repeated in order to continue developing the narrative, as in (40).

(40) *Pije o pulakis na pai na pari fajito ja ta pulakia tu ke itane **ena gataki**. **To gataki** kitakse ti folia me ta pulakia mikrutsika*[13].

'The little bird went to go to take food for his little birds and there was **a kitten**. **The kitten** looked at the nest with the little birds.'

<div align="right">(S153, HS, age: 57, intermediate)</div>

(ii) LS preceded by NS:

LS in TC were sometimes preceded by a clause with a corefering NS (19 out of 162). In example (41), the previous explicit mention of the referent was not very close to the LS in question. This may imply that the speaker uses LS because s/he may be sensitive to the possibility of misunderstanding by the listener (Hendriks et al. 2014).

(41) *Ena aloyataki etrexe na ði to filo tu- ti filenaða tu ajelaða. Ala ixane vali enan ftaxti, tote ðen boruse na pa na tin ði. Ala **to aloyataki** pije na piðiksi to fraxti.*

'A little horse was running to see his friend the cow. But [they] had put a fence, so he could not go to see her. But **the little horse** went to jump the fence.'

<div align="right">(S26, HS, age: 16, near-native)</div>

(iii) LS forming equivalents:

Some LS in TC were lexical units embodying the meaning of the subject of the previous clause (9 out of 162). The notional subject was identical although the grammatical subject was different. In example (42), the subject *sus nuevos amigos* ('his new friends') are *el pájaro y el toro* ('the bird and the bull'). This was made for stylistic purposes, to give a supplementary meaning.

(42) *Entonces **sus nuevos amigos** lo vieron en problemas. **El pájaro y el toro** tomaron su set de primeros auxilios y asistieron al caballito con su pata herida.*

'Then **his new friends** saw him in trouble. **The bird and the bull** took their first-aid kit and helped the little horse with his wounded leg.'

<div align="right">(S85, Spanish, age: 37)</div>

---

[13] The adjective-noun order here is transferred from Spanish. This is ungrammatical in Greek.

(iv) LS in rephrasing:

Some LS in TC were repeated for rephrasing purposes (6 out of 162). This occurred as a form of self-repair or reflecting stylistic choices, as in (43).

(43)   *Se mia stigmi efije **i peristera**, petakse **i peristera** na vri fai ja ta pulakia tis.*

'At one moment the dove left, the dove flew to find food for her babies.'

(S32, IMM, age: 84, LoR: 29)

(v) LS in focus:

In Greek monolinguals, 4 cases involved a clause containing the LS which was focused, as in (44) in which the focus is on the predicate *ine kala* ('is well').

(44)   ***To aloyaki** ine zondano ke lipon, ne, **to aloyaki** <u>ine kala</u>.*

'**The little horse** is alive and so, yes, **the little horse** is well.'

(S99, Greek, age: 36)

(b) Redundant LS

Overall, 22.83% of LS in TC (N=37) were redundant (infelicitous) since their omission would not affect the interpretation or style of the utterance (examples 45-47).

(45)   ***El caballo** da un brinco muy grande y quiere saltar la cerca, pero desgraciadamente **el caballo** se tropieza y cae de espalda.*

'**The horse** jumps very high and wants to jump the fence but unfortunately **the horse** stumbles and falls on his back.'

(S67, Spanish, age: 59)

(46)   *[...] **to aloyaki** etrexe anemelo ston kambo. Ke **to aloyaki** sinandise ðio filus tu.*

'[...] **the little horse** was running in the field. And **the little horse** met two friends of his.'

(S143, IMM, age: 22, LoR: 3)

(47) *[...] **to yato** efije ke otan **to yato** efije, ksanaerthi to meyalo puli me fajito ja ta pulakia.*

'[...] **the cat** left and when **the cat** left, the big bird returned with food for the little birds.'

(S47, L2, age: 30, intermediate)

Interim summary

Table 6.16 displays the number and total percentage of non-redundant and redundant LS in TC. No statistical differences were found between the groups. Although HS and L2 groups demonstrated an increased use of LS in TC, which was significantly different compared to the other groups (§6.3.1.1), the qualitative approach (§6.3.1.2) revealed that all groups were similar regarding LS redundancy.

Table 6.16. Number and total percentage of (non-)redundant lexical subjects in TC

Lexical subjects in TC						
Non-redundant LS	**Spanish**	**Greek**	**IMM**	**HS**	**L2**	**Total**
Following presentational focus	17	9	24	12	25	87
Preceded by corefering NS	3	1	2	8	5	19
Forming equivalents	2	3	1	1	2	9
Rephrasing	0	1	2	0	3	6
Focus	0	4	0	0	0	4
Total	22	18	29	21	35	125
%	*78.6*	*78.3*	*74.4*	*80.8*	*76.1*	*77.2*
Redundant LS						
Total	6	5	10	5	11	37
%	*21.4*	*21.7*	*25.6*	*19.2*	*23.9*	*22.8*
Total	28	23	39	26	46	162
%	*100*	*100*	*100*	*100*	*100*	*100*

(Over)use of overt subject pronouns in topic continuity

OSP in TC are infelicitous unless focused or bearing emphatic/contrastive stress. Only 11 OSP (personal/demonstrative) emerged in TC (Spanish: 2, Greek monolinguals: 1, IMM: 2, HS: 1 and L2ers: 5), which was 0.42% of TC contexts. All occurrences, excepting those of the L2 group, involved contrastiveness or emphasis, which was expressed with personal pronouns used anaphorically (48, 49, 52) or deictically (53) or with demonstrative pronouns in anaphoric position (50, 51, 54). In (50) and (51) the presence of the focaliser particle *ki (ke)* 'and/too/neither' clearly conveys focus[14].

(48) *Ella recién había tenido estos pajaritos y tenía que alimentarlos, porque así es en la familia de los pájaros. Dejan a las crías en el nido y **ellos** salen a buscar el alimento y lo traen en su piquito para sus bebés.*

'She recently had these little birds and had to feed them, because this is the way in the birds' family. [They] leave the young in the nest and **they** leave to look for food and [they] bring it in their beak for their babies.'

(S69, Spanish, age: 70)

(49) *El caballo intentó pasar sobre la cerca, pero al parecer corrió mal y se tropezó en la cerca, entonces la desestabilizó y la rompió. Y a la vez **él** se pegó muy fuerte y cayó al piso.*

'The horse tried to pass over the fence but seemingly [he] ran badly and [he] tripped over the fence, then [he] destabilised it and [he] broke it. And at the same time **he** got hurt very badly and [he] fell on the ground.'

(S73, Spanish, age: 28)

(50) *Arxizi na skarfaloni xoris **ute ki ekini** na proseksi oti akrivos apo piso tis vriskete o kalos angelos ton mikruliðon.*

'[She] starts to climb without **[she neither]** noticing that right behind her there is the good angel of the little ones.'

(S93, Greek, age: 52)

---

[14] See Holton et al. (2012); Chatzikyriakidis, Michelioudakis & Spathas (2015).

(51) *[...] ipirxe ena aloyaki sto ipethro [...] iðe mia ajelaða [...], eki erotheftike nomizo tin ajelaða [...]. I ajelaða kathotane ke evlepe.* **Ki ekini** *eroteftike ton alogaki.*

'[...] there was a little horse in the field [...] [he] saw a cow [...], there [he] fell in love I think with the cow [...]. The cow was sitting and watching. **She too** fell in love with the little horse.'

(S32, IMM, age 84, LoR: 29)

(52) *Ta kitai lipon i γata apo kato apo to ðendro ke prospathi, epiði mallon pinai, na anevi. Tin ora omos pu prospathi* **afti** *na anevi pano ke na fai ta pulia, jati lipi i mama, pernai apo 'ki pera enas skilos.*

'The cat looks at them from under the tree and [she] tries, because probably [she] is hungry, to go up. At the time though that **she** tries to go up and eat the birds, because the mother is not there, a dog passes by.'

(S59, IMM, age: 45, LoR: 3)

(53) *Ena poli oreo alogo [...] theli na piðisi, 'pero' epese, epese* **afto**. *Ke to oreo to puli, to pulaki, efere mia valitsa me ta prota voithia.*

'A very nice horse [...] wants to jump, but [he] fell, **he** fell. And the nice bird, the little bird, brought a suitcase with the first aids.'

(S150, HS, age: 54, intermediate)

(54) *Lipon, iparxi ena ðendro pu ine ena pulaki me ta tria- me ala tria pulakia.* **Afti** *ine i mama ke- ke iparxi mia γata pu kitai ta mikra pulakia.*

'So, there is a tree where there is a little bird with the three- with three other little birds. **She** is the mum and- and there is a cat who looks at the little birds.'

(S16, L2, age: 36, advanced)

There were 5 OSP in TC in the L2ers' production data. However, 4 out of 5 cases were produced by one low-proficiency speaker (S64). In his oral performance, *aftos* in (55) is possibly emphasised, while the OSP used in (56)-(58) can be considered redundant, thus infelicitous, since they are pleonastic.

(55)   *Tote **o aloyos** ðen mbori na pai eki, ala **aftos** piðai to fraxtis.*

‘Then the horse cannot go there, but **he** jumps the fence.’

(56)   ***O aloyos** theli na pai pu ine i ajelaða, ala otan **aftos** piðai, kani ena lathos ke ‘cae’, siɣnomi, ke pefti.*

‘The horse wants to go where the cow is, but when **he** jumps, [he] makes a mistake and [he] falls.’

(57)   *[...] **enas skilos** erthi- erthi sti skini ki **aftos** ðaɣoni ton ɣato.*

‘[...] a dog comes in the scene and **he** bites the cat.’

(58)   *I ɣata kitai- kitai tin folia ke nomizo **afto** theli na pai eki, jati iparxi mikres pulies.*

‘The cat looks at the nest and I think **she** wants to go there because there are little birds.’

(S64, L2, age 31, basic)

From the above-quoted examples involving OSP in TC contexts, it can be observed that no overgeneralization of the scope of the OSP was found in bilinguals. An exception is one individual case in the L2 group, which can be regarded as an outlier.

6.3.1.3 Summary and discussion

In the contexts of TC, it was observed that monolingual and bilingual groups of speakers expressed same reference by using predominantly NS. Greek monolinguals produced significantly more NS in TC than the Spanish, the HS and the L2 group in matrix clauses. LS were also used in TC by all groups but were significantly more in HS and especially in L2ers when compared to the other groups. Scrutinising qualitatively the particular contexts of TC with LS, it was observed that most occurrences did not involve redundant LS. Thus, although HS and L2ers overused LS in TC, their performance was similar to Greek monolingual and immigrants respecting redundancy. It was thus noted that redundancy is context-dependent and LS in TC do not a piori imply infelicity.

171

Overall, despite statistical differences in the relative frequency of subjects used in TC (§6.3.1.1), qualitatively all groups behaved similarly (§6.3.1.2), revealing a generally homogeneous behaviour on subjects' distribution in TC. Crucially, there was no overuse of OSP in TC contexts in the bilingual performance. The individual results of one L2 speaker who produced pragmatically inappropriate OSP in this context are not generalisable.

Without regard to particular details, performance in production of TC could be seen as conforming to postulations of accessibility accounts (e.g. Ariel 1990; Gundel et al. 1993; Carminati 2002). Accordingly, null referential subjects display biases towards establishing coreference with highly salient antecedents, such as subjects. The scarcity of OSP contexts indicates no overuse by bilinguals. These findings are further discussed in §6.6.

6.3.2 Topic Shift

6.3.2.1 Quantitative analysis

The preferred way of marking TS in narratives was with LS in all groups of speakers (63.54%-76.77%) (Figure 6.11, Table 6.17). Greek monolinguals used LS in TS more often than Spanish monolinguals (74.71% vs 66.52%) and this difference was statistically significant [Pearson χ^2 (1, N=490)=3.96, p=0.047]. In Spanish, OSP were used significantly more than in monolingual Greek (8.15% vs 1.95%) [Pearson χ^2 (1, N=490)=10.11, p=0.001]. Spanish monolinguals used OSP considerably more than all the Greek-speaking groups (8.15% vs a maximum of 1.95%). No significant differences were found between the Greek-speaking groups regarding OSP.

Greek monolinguals produced significantly more LS in TS than HS (74.71% vs 63.54%) [Pearson χ^2 (1, N=449)=6.51, p=0.011] and did not statistically differ from immigrants and L2ers in LS use. The groups of immigrants and HS were similar regarding LS use. Statistical significance was reached in comparing LS rates in TS between L2 and HS (76.77% vs 63.54%) [Pearson χ^2 (1, N=390)=8.15, p=0.004].

All groups of speakers employed NS in TS (21.21%-34.9%). The relative frequency of NS was higher than that of OSP in this context. HS exhibited a significantly higher rate of NS in TS (34.9%) compared to Greek monolinguals (23.35%) [Pearson χ^2 (1, N=449)=7.22, p=0.007] and L2ers (21.21%) [Pearson χ^2 (1, N=390)=9.06, p=0.003]. The difference between immigrants (29.04%) and HS was not statistically significant. Additionally, the L2 group significantly differed from the immigrants in the rates of NS in TS (21.21% vs 29.04%) [Pearson χ^2 (1, N=594)=4.16, p=0.041], although they did not differ in LS rates [Pearson χ^2 (1, N=594)=3.05, p=0.080].
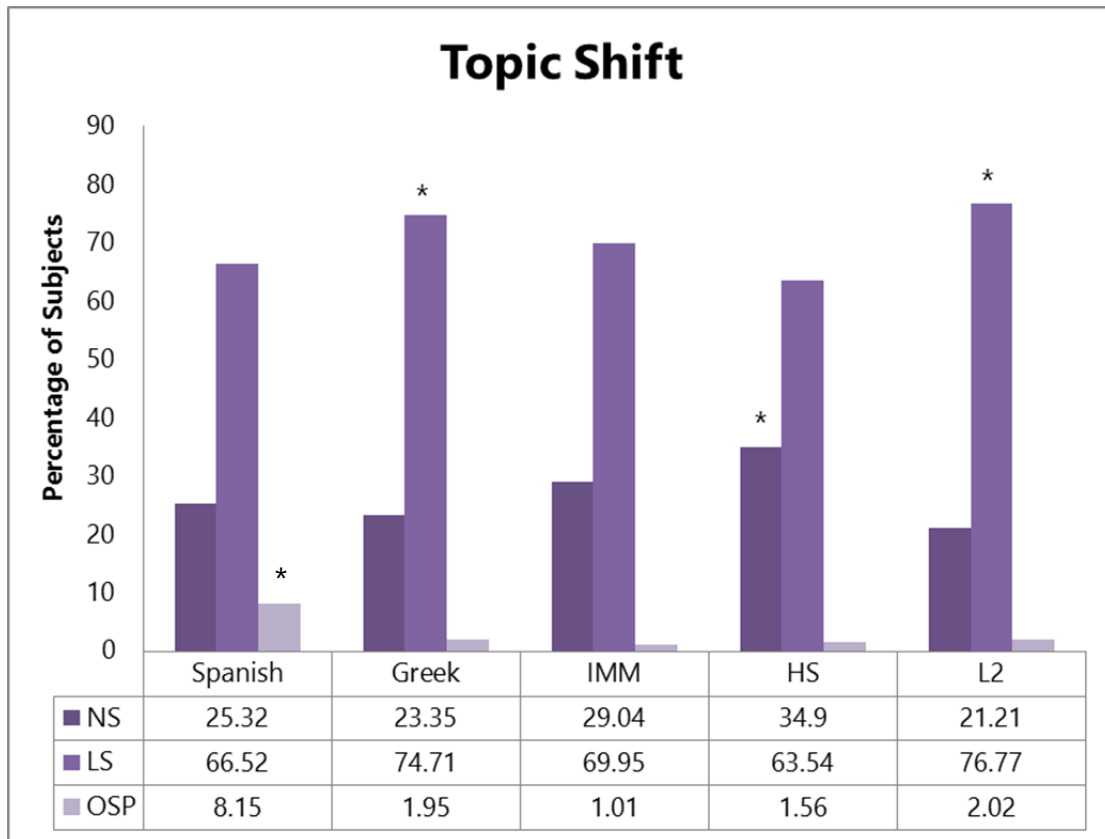
Figure 6.11. Use of subjects in TS contexts

Table 6.17. Use of subjects in TS contexts

	Spanish	Greek	IMM	HS	L2	Total
Null Pronoun	59	60	115	67	42	343
	25.32	*23.35*	*29.04*	*34.9*	*21.21*	*26.88*
Lexical Subject	155	192	277	122	152	898
	66.52	*74.71*	*69.95*	*63.54*	*76.77*	*70.38*
Overt Pronoun	19	5	4	3	4	35
	8.15	*1.95*	*1.01*	*1.56*	*2.02*	*2.74*
Total	233	257	396	192	198	1,276
	100	*100*	*100*	*100*	*100*	*100*

Considering only the matrix clauses (Figure 6.12, Table 6.18), the percentage of OSP was significantly higher in the Spanish group relative to Greek monolinguals [Pearson χ^2 (1, N=373)=7.10, Fisher's p=0.008]. Significant differences were observed in NS in TS (a) between HS and monolingual Greek [Pearson χ^2 (1, N=352)=4.50, p=0.034] and (b) between HS and L2 group [Pearson χ^2 (1, N=305)=5.01, p=0.025]. HS produced significantly more NS than Greek monolinguals and L2ers.

Topic Shift: Matrix Clauses

	Spanish	Greek	IMM	HS	L2
■ NS	17.86	19.61	24.42	29.25	18.35
■ LS	74.4	78.43	74.59	69.39	79.75
■ OSP	7.74	1.96	0.99	1.36	1.9
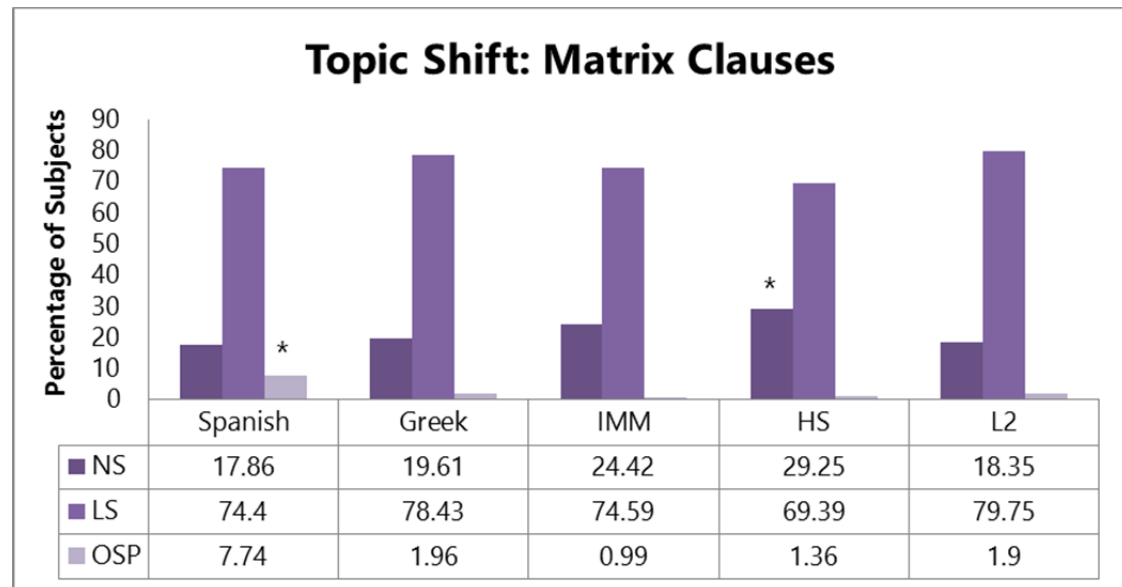
Figure 6.12. Use of subjects in TS contexts in matrix clauses

Table 6.18. Use of subjects in TS contexts in matrix clauses

	Spanish	Greek	IMM	HS	L2	Total
Null pronoun	30	40	74	43	29	216
	_17.86_	_19.51_	_24.42_	_29.25_	_18.35_	_22.02_
Lexical subject	125	161	226	102	126	740
	_74.4_	_78.54_	_74.59_	_69.39_	_79.75_	_75.43_
Overt pronoun	13	4	3	2	3	25
	_7.74_	_1.95_	_0.99_	_1.36_	_1.9_	_2.55_
Total	168	205	303	147	158	981
	_100_	_100_	_100_	_100_	_100_	_100_

In embedded clauses, LS were mainly used in TS by all groups excepting HS (Figure 6.13, Table 6.19). LS were used less while NS were used more than in matrix clauses. Spanish and Greek monolinguals were not significantly different in OSP frequencies. No statistical differences were observed in the relative frequencies of subjects within and between groups, but the L2 group tended to employ much more LS than NS in embeddings.
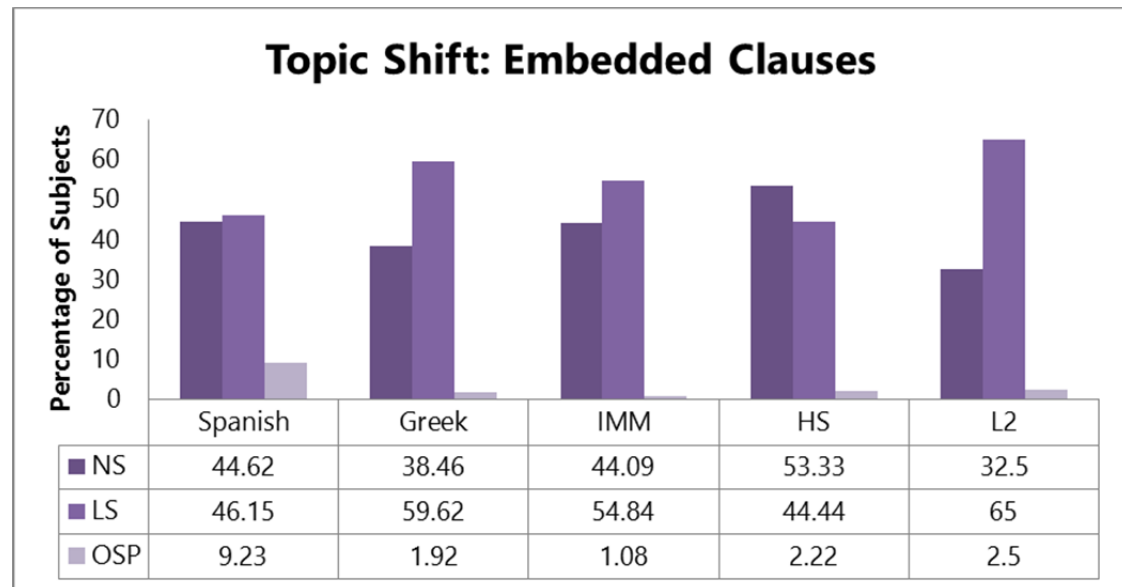
Figure 6.13. Use of subjects in TS contexts in embedded clauses

Table 6.19. Use of subjects in TS contexts in embedded clauses

	Spanish	Greek	IMM	HS	L2	Total
Null pronoun	29	20	41	24	13	127
	44.62	*38.46*	*44.09*	*53.33*	*32.5*	*43.05*
Lexical subject	30	31	51	20	26	158
	46.15	*59.62*	*54.84*	*44.44*	*65*	*53.56*
Overt pronoun	6	1	1	1	1	10
	9.23	*1.92*	*1.08*	*2.22*	*2.5*	*3.39*
Total	65	52	93	45	40	295
	100	*100*	*100*	*100*	*100*	*100*

6.3.2.2 Qualitative analysis

All groups of speakers produced more LS than other categories of subjects in TS. All groups also produced NS in TS. A closer look at the contexts of TS offers an insightful view on how it is (in)felicitously expressed regarding subject reference.

Use of lexical subjects in topic shift

All groups of speakers predominantly used LS in of TS, as in (59) and (60).

(59) *Finalmente **el gato** baja del árbol, porque **el perro** lo sale persiguiendo y **el pájaro adulto** se acerca al nido.*

'Finally **the cat** gets down the tree, because **the dog** starts chasing him and **the adult bird** approaches the nest.'

(S66, Spanish, age: 29)

(60) *Erxete **o skilos** ke erxete **to pulaki** me to fai ja ta mikra ki **o skilos** travai tin ura ke pefti **to gato** kato.*

'**The dog** comes and **the little bird** comes with the food for the little ones and **the dog** pulls the tail and **the cat** falls down.'

(S45, HS, age: 65, near-native)

Use of overt subject pronouns in topic shift

TS can be encoded with OSP in referent reintroduction. There were 35 cases of OSP in TS in the corpus, which was 2.74% of the TS contexts. More than half of these cases (N=19) were produced by Spanish speakers and the remaining (N=16) by the four Greek-speaking groups (monolinguals: 5, IMM: 4, HS: 3, L2: 4). All groups used OSP sparingly (see §6.3.2.1).

In Spanish, 11 out of 20 speakers used 19 OSP in TS (8.15%) with each speaker producing a maximum of 3 OSP in this context (examples 61-69). While 16 OSP involved personal pronouns, there were 3 cases of demonstratives (e.g. 65, 66).

(61) *[...] tras él hay un perro que lo jala de la cola y consigue que **él** no llegue y no pueda trepar hasta el nido.*

'[...] behind him there is a dog who pulls him from the tail and manages that **he** does not reach and is not able to climb to the nest.'

(S67, Spanish, age: 59)

(62) *Luego la vaca se aleja y **él** corre para acercarse a la vaca.*

'Then the cow walks away and **he** runs to come near the cow.'

(S70, Spanish, age: 39)

(63) *[...] le ladró muy fuerte para que **él** se ahuyentara.*

'[...] he barked at him very loudly so that **he** would be chased away.'

(S78, Spanish, age: 45)

(64) *[...] se veían casi todos los días cada vez que **él** estaba en el prado.*

'[...] they saw each other almost every day every time **he** was in the meadow.'

(S84, Spanish, age: 39)

(65) *La mamá pájaro estaba con sus polluelos en el nido y **estos** gritaban de hambre. Entonces **ella** voló a buscar comida y los dejó solitos.*

'The mother bird was with her chicks in the nest and **they** were crying because of hunger. Then **she** flew to find food and left them alone.'

(S82, Spanish, age: 39)

(66) *El caballito, al saltar la cerca, dio un mal paso y **ésta** se quebró.*

'The little horse, when jumping the fence, took a bad step and **it** broke.'

(S75, Spanish, age: 46)

In 9 out of 19 OSP in TS in Spanish (47.4%), the verbs were in third-person singular past imperfect (7 cases, e.g. 64) or subjunctive (2 cases: 61, 63) involving forms with ambiguous morphology in terms of person features (see §2.3.7). In (63), the verb form is actually not ambiguous due to the presence of the reflexive *se*, which indicates

178

third person turning the verb in unaccusative. Subjunctive in the embedded (complement) clause in (63) marks obligatory disjoint reference between the OSP *él* and the subject on which it depends (*el perro*). Although there was TS, the OSP was not crucial since the verb form in this type of embedding disallows same reference (TC) interpretation and there were only two possible antecedents. In 6 out of 9 OSP in TS involving ambiguous verbal morphology, the OSP was not required for disambiguation. Overall, from 19 OSP in TS, 6 were required to disambiguate reference (e.g. 66) or to avoid temporary ambiguity (e.g. 62). In the remaining 13 cases (e.g. 63, 64), the presence of OSP was not vital for referential disambiguation. In sum, even in TS most of OSP in Spanish i.e. 13 out of 19 (11 out of 16 personal and 2 out of 3 demonstrative pronouns) would not cause ambiguity if omitted.

In monolingual Greek, 4 out of 20 of speakers used 5 OSP in TS (1.95%) (examples 67-71). In (67) and (68), the OSP are focused by the focus-associated operator *ke* ('too'). All 5 OSP were required to disambiguate reference and/or to convey focus.

(67) *I ajelaða ton kitakse, tin kitakse **ki aftos** ke pire tin apofasi na piðiksi ton fraxti.*

'The cow looked at him, **he** looked at her **too** and decided to jump the fence.'

(S101, Greek, age: 42)

(68) *I ajelaða lipate, to puli meni kataplikto, ala omos **ke afta** fanikan xrisima.*

'The cow was sorry, the bird was surprised, but **they too** were useful.'

(S103, Greek, age: 64)

(69) *Ti γrapose liγo prin **ekini** ftasi sti folia me ta tria pulakia.*

'He grabbed her just before **she** reached the nest with the three little birds.'

(S101, Greek, age: 42)

(70) *I manula exi petaksi makria ke **afti** vriski tin efkeria ke aneveni sto ðendro.*

'The mammy had flown away and **she** [the cat] finds the opportunity to climb up the tree.'

(S93, Greek, age: 52)

(71) *Ixe erthi i mama m' ena skulikaki sto stoma na ta taisi. **Afta** den to piran xambari.*

'The mother had come with a small worm in the mouth to feed them. **They** did not realise anything.'

<div align="right">(S95, Greek, age: 48)</div>

In the immigrants' group, 3 out of 35 speakers produced a total of 4 OSP in TS (1.01%), shown in (72)-(75). It can be argued that in two cases, i.e. (73) and (75), the presence of the OSP was vital for reference disambiguation.

(72) *Itane mia kotula, ixe tria avyulakia, ta zestane ke vyikan tria pulakia. Tora **afti** prepi na pai na feri fai na taisi ta peðia tis.*

'There was a little hen, [she] had three little eggs, [she] warmed them and three little birds went out. Now **she** has to go to bring food to feed her young.'

(73) *Tora molis efije i mitera, **afto** pije n' anevi ja na fai ta kaimena ta pulakia.*

'Now when the mother left, **she** started going up to eat the poor little birds.'

<div align="right">(S09, IMM, age: 80, LoR: 54)</div>

(74) *Ala ta pedia tis leyane «mama, mama, pinao». Opote **afti** pire tin apofasi na pai na vri fajito ja ta peðia tis.*

'But the young told her "mum, mum, [I] am hungry". So **she** took the decision to go to find food for her young.'

<div align="right">(S143, IMM, age: 22, LoR: 3)</div>

(75) *[…] ki etsi ðagose ti yata apo tin ura liyo prin **afti** kataferi na ftasi sti folia.*

'[…] so [he] bit the cat from the tail just before **she** managed to reach the nest.'

<div align="right">(S52, IMM, age: 39, LoR: 4)</div>

In the HS group, 3 out of 21 speakers used 3 OSP in TS (1.56%) shown in (76)-(78). The OSP was not optional in the discourse context in (77) and (78).

(76) *I mama pulaki ðen iðe afto pu ekane o skilos ala **afti** jirise me fajito ja ta tria pulakia tis.*

'The mother bird did not see what the dog did but **she** came back with food for her three little birds.'

(S23, HS, age: 29, intermediate)

(77) *Ke tin kalokitusa ke **ekini** episis tin- tu kalokitusa [...].*

'And [he] stared at her and **she** also stared at her- him [...].'

(S30, HS, age: 49, basic)

(78) *Ena aloyaki pu trexi mes ton kambo vlepi ena klisimo me sirmata pu xorizi to meros pu **aftos** ine me ena- me mia ali farma, stin opia exi mia ajelaða.*

'A little horse that runs in the field sees a closure with wires which divides the place where **he** is from a- from another farm, in which there is a cow.'

(S36, HS, age: 57, advanced)

In the L2 group, 2 out of 20 speakers used 4 OSP in TS (1.56%), with each speaker producing 2 OSP, as in (79). Both speakers were of basic proficiency presenting disfluency and some morphological errors, so OSP facilitated reference interpretation.

(79) *Iparxi luluði episis, luluðia, ke **aftos** kitai ena luluði ke ine ena ajelaða. Nomizo **afto** theli na pai pu i ajelaða ine.*

'There is a flower also, flowers, and **he** looks at a flower and there is a cow. I think **he** wants to go where the cow is.'

(S64, L2, age 31, basic)

A comparison of OSP use in TS between Greek and Spanish shows that OSP in monolingual Greek and in most cases in the bilingual performance were typically used when required for ambiguity avoidance. By contrast, in Spanish more than half OSP were not absolutely necessary for reference disambiguation.

6.3.2.3 Ambiguity

Quantitative analysis

Due to the potential ambiguity of NS in TS regarding the subject referent of the clause, occurrences of NS in TS were examined in more detail. The participants were instructed to produce narratives for a hypothetical listener who could not see the pictures. They were thus expected to be explicit and to produce unambiguous utterances since any under-explicit clause could hinder communication.

NS produced in TS were marked as ambiguous or non-ambiguous after examining their context. No potential ambiguity was defined along the lines of Shin (2014: 317), i.e. 'reference could be clearly established because the referent was identified (anywhere) in the preceding discourse and there were no pragmatically viable competing referents'.

Figure 6.14 and Table 6.20 show the numbers and percentages of ambiguous and non-ambiguous NS in TS. The majority of NS in TS were not ambiguous, with a range of 77.61% to 98.31% in all groups. The monolinguals showed very small rates of ambiguous NS with only 1 case in Spanish (1.69%) and 2 cases in Greek (3.33%). The bilinguals manifested greater percentages of ambiguous NS. The immigrants produced 17 cases of ambiguity (14.78%), followed by L2ers with 9 cases (21.43%) and HS with 15 cases (22.39%).

Pearson Chi-square tests revealed significant differences regarding ambiguity rates between Greek monolinguals (3.33%) and immigrants (14.78%) [Pearson χ^2 (1, N=175)=5.34, Fisher's p=0.021], L2ers (21.43%) [Pearson χ^2 (1, N=102)=62.04, Fisher's p<0.001] and HS (22.39%) [Pearson χ^2 (1, N=127)=71.45, Fisher's p<0.001]. No significant differences were detected between the groups of bilinguals.
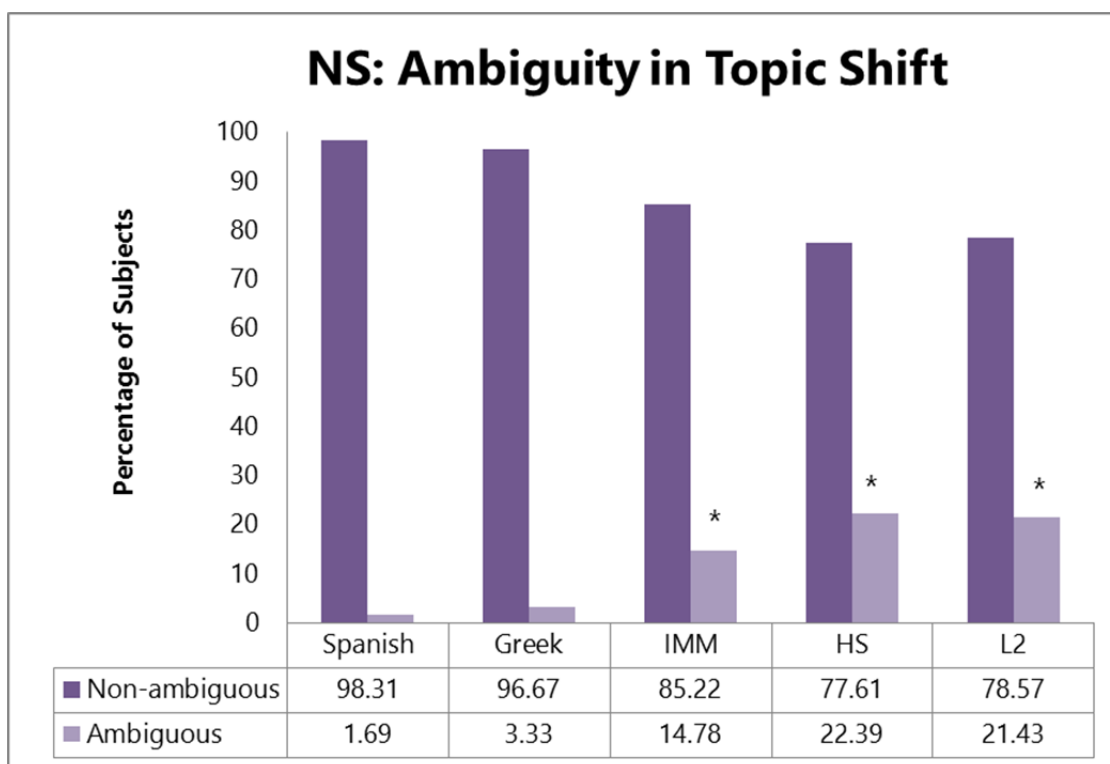
Figure 6.14. Ambiguous and non-ambiguous NS in TS

Table 6.20. Ambiguous and non-ambiguous NS in TS

	Spanish	Greek	IMM	HS	L2	Total
Non-ambiguous	58	58	98	52	33	299
	98.31	*96.67*	*85.22*	*77.61*	*78.57*	*87.17*
Ambiguous	1	2	17	15	9	44
	1.69	*3.33*	*14.78*	*22.39*	*21.43*	*12.83*
Total	59	60	115	67	42	343
	100	*100*	*100*	*100*	*100*	*100*

In order to explore the factors that affect occurrence of ambiguity in production, the variable of Age at testing and Proficiency were examined. The relationship between Age and Ambiguity was analysed performing a non-parametric t-test (Wilcoxon-Mann-Whitney). Results showed a significant association (N=343, z=-3.456, p<0.001) between the variables indicating that more cases of ambiguity were produced by participants of older ages. Table 6.21 demonstrates the mean ages and SD of speakers who produced non-ambiguous and ambiguous cases involving NS in TS.

Table 6.21. Age of speakers who produced non-ambiguous and ambiguous NS in TS

	Non-ambiguous NS					Ambiguous NS				
	Obs	Mean	SD	Min	Max	Obs	Mean	SD	Min	Max
Age	299	48.36	17.76	16	86	44	59.11	19.84	16	84

The mean age of speakers who produced non-ambiguous NS in TS was 48.4 years, while the mean age of those who produced ambiguous NS in TS was 59.1 years. It is thus observed that ambiguous NS were generally produced by older participants.

Proficiency was not found to be significant in relation to ambiguity in the groups of HS and L2. As shown in Table 6.22, 24 ambiguous cases of NS in TS contexts were produced by HS and L2ers in the four proficiency levels and no clear pattern emerged in these occurrences with regard to proficiency. In fact, half of the ambiguity cases (N=12) were produced by basic and intermediate level speakers and the other half (N=12) by advanced and near-native speakers.

Table 6.22. Proficiency of HS and L2 in non-ambiguous and ambiguous NS in TS

	Basic	Intermediate	Advanced	Near-native	Total
Non-ambiguous	14	17	34	20	85
	63.64	*80.95*	*87.18*	*74.07*	*77.98*
Ambiguous	8	4	5	7	24
	36.36	*19.05*	*12.82*	*25.93*	*22.02*
Total	22	21	39	27	109
	100	*100*	*100*	*100*	*100*

It is known, however, that proficiency correlates with accuracy in agreement morphology (e.g. Montrul & Rodríguez Louro 2006; see also §6.3, Table 6.11). This will be clarified subsequently in the qualitative analysis. The qualitative analysis of the cases of NS in TS focusing on cases of ambiguity was necessary and in order to reveal aspects which do not emerge through quantitative analyses.

Qualitative analysis

NS in TS are not infelicitous if the subject referent is salient enough and can be identified through cues in the discourse and inferences. If the context contains a NS in TS but not sufficient cues to identify its referent due to presence of competing referent(s), the NS is ambiguous (infelicitous) and may lead to miscommunication. The majority of NS in TS were found to be non-ambiguous (N=299, 87.17%), whereas a number of NS were marked as ambiguous (N=44, 12.83%). The ambiguity of NS was not always due to a missing overt subject, as will be illustrated shortly.

(a) Non-ambiguous NS

All groups used unambiguous NS in TS. Non-ambiguity was due to morphological, semantic or contextual cues, which removed uncertainty regarding the subject referent. Table 6.23 shows the numbers and percentages of unambiguous NS in TS (from Table 6.20) and the respective speakers. Examples are offered in (80)-(83).

Table 6.23. Use of unambiguous NS in TS contexts in all groups

Group	Unambiguous NS in TS		Speakers using unambiguous NS in TS		Maximum N of unambiguous NS in TS per speaker
	N	%	N	%	N
Spanish	58/59	*98.31*	19/20	*95*	6
Greek	58/60	*96.67*	18/20	*90*	8
IMM	98/115	*85.22*	34/35	*97.2*	9
HS	52/67	*77.61*	16/21	*76.2*	7
L2	33/42	*78.57*	17/20	*85*	6

(80) *[...] la vaca está sanándole una herida. Seguramente Ø se hirió y entonces Ø lo está sanando.*

'[...] the cow is healing him the wound. Probably **[he]** was hurt and then **[she]** is healing him.'

(S68, Spanish, age: 28)

185

(81)   *Tin ora pu i gata prospathi n' anevi pano sti folia, Ø tin arpazi apo tin ura ke tin*

   katevazi kato.

   'At the time when the cat tries to climb up to the nest, **[he]** grabs her form the

   tail and [he] draws her down.'

<div align="right">(S105, Greek, age: 70)</div>

(82)   *Ena aloyataki etrexe na ði to filo tu- ti filenaða tu ajelaða. Ala Ø ixane vali enan*

   ftaxti tote Ø ðen mboruse na pa na tin ði.

   'A little horse was running to see his friend- his girlfriend the cow. But **[they]**

   had put a fence and **[he]** so he could not go to see her.'

<div align="right">(S26, HS, age: 16, near-native)</div>

(83)   *Ena skilo piani ti yata apo tin ura Ø ja na min piasi tin folia ton pulion.*

   'A dog grabs the cat from the tail so that **[she]** does not catch the birds' nest.'

<div align="right">(S33, L2, age: 43, near-native)</div>

(b) Ambiguous NS

Ambiguity involving NS in TS contexts emerged as the result of competing referents. The bilingual groups produced significantly more ambiguous NS in TS contexts than monolinguals. Table 6.24 displays the numbers and percentages of ambiguous NS in TS (from Table 6.20) and the respective speakers.

Table 6.24. Use of ambiguous NS in TS contexts in all groups

Group	Ambiguous NS in TS		Speakers using ambiguous NS in TS		Maximum N of ambiguous NS in TS per speaker
	N	**%**	**N**	**%**	**N**
Spanish	1/59	*1.69*	1/20	*5*	1
Greek	2/60	*3.33*	2/20	*10*	1
IMM	17/115	*14.78*	9/35	*25.7*	3
HS	15/67	*22.39*	10/21	*47.6*	3
L2	9/42	*21.43*	7/20	*35*	2

The ambiguous NS in TS found in the production data involved cases of referent reintroduction in subject position. In some cases, such reintroduction would require an overt subject (lexical or pronominal) taking into account the listener's perspective (see Gundel et al. 1993; Liceras et al. 2010; Hendriks et al. 2014). In other cases, an overt subject was not obligatory to prevent ambiguity, since ambiguity was triggered by other problems in the expression of the utterance. If the TS is not perceived due to use of NS and absence (or incorrect use) of other appropriate cues, the referent may be misunderstood as erroneously referring to the subject of the previous clause or to some other referent. In other words, the resolution may result in a TC interpretation or ambiguity, hence miscommunication and/or processing cost.

In the cases shown next, NS interpretation was temporarily or permanently biased towards TC (erroneous interpretation) or it caused ambiguity perceived by the listener of the narrative (incomplete interpretation). Such cases of infelicity due to underspecified constructions were however unintentionally produced by the speakers, who did have a specific meaning in mind, yet it could not be conveyed effectively. In this sense, the ambiguous clauses may be regarded as vague rather than ambiguous (see Wasow 2015). Nonetheless, since vagueness (lack of preciseness) entails ambiguity, i.e. the quality of being open to more than one interpretation, the term ambiguity is employed in the present study.

Ambiguity was classified into two main types: (i) temporary and (ii) full ambiguity.

(i) Temporary or local ambiguity is generally resolved through the continuation of the utterance once the listener has mentally processed it. Disambiguation takes place at a certain point in the communication process by means of morphological or semantic cues that may appear later in the discourse. Temporary ambiguity may be also resolved with the help of the context, world knowledge and human inferential abilities, assuming cooperation by the listener. Correct referent identification sometimes may stem from metalinguistic comments given by the speaker to explicitly specify the intended meaning (Wasow 2015; Winter-Froemel & Angelika Zirker 2015).

(ii) Full ambiguity causes more than one possible interpretation of the referent in the utterance. Since neither of the possible interpretations can be ruled out, it may thus confuse or misguide the listener. Full ambiguity often remains unresolved (Wasow 2015). In the present study full ambiguity is related to either morphological misconstructions or to genuine ambiguity caused by the lack of explicit subject.

Table 6.25 displays the number of cases of ambiguity concerning NS in TS per group of speaker according to ambiguity type.

Table 6.25. Ambiguous NS in TS per type of ambiguity and group of speakers

Ambiguity	N					
	Spanish	Greek	IMM	HS	L2	Total
Temporary ambiguity	1	1	7	7	0	16
Full ambiguity						
Morphological ambiguity	0	0	2	6	2	10
Genuine ambiguity	0	1	8	2	7	18
Total	1	2	17	15	9	44

(i)Temporary ambiguity

Cases of temporary ambiguity, i.e. ambiguity which is most likely to be properly resolved, were triggered by a NS and eventually resolved through morphology, repairs, contextual or pragmatic cues. Such cases emerged in the production of immigrants and HS with 7 cases found in each group, while no such cases were found in the L2 group (example 84).

(84) *Anevene to ðendro ala enas skilos iðe ton γato ke iksere oti ithele na fai ta pulakia. Tote tu Ø² **ðagose** tin ura ke to travikse Ø² ja **na fiji**.*

'[He] was climbing up the tree but a dog saw the cat and [he] knew that [he] wanted to eat the little birds. Then **[he²]** bit his tail and pulled him so that **[he²]** goes away.'

(S26, HS, age: 16, near-native)

In (81), the subject referent of the verb *ðagose* ('bit') may apparently be the cat and not the dog, the latter being the correct/intended referent. Thus, the context may seem to be TC and not TS. This is resolved pragmatically through world knowledge, i.e. dogs in general are more likely to bite animals than cats. The cat is indeed a possible antecedent but a less plausible one than the dog. Hence, the ambiguity is most probably resolved. The use of OSP would not necessarily help disambiguation since the two referents are masculine. The resolution of the NS in *na fiji* ('to leave') is also ambiguous since it depends on the resolution of the NS in *ðagose* ('bit'), but it is eventually resolved referencing the cat. Due to the fact that temporary ambiguity usually does not remain unresolved, it is considered to be less problematic than full ambiguity, which is the focus of the next subsection.

(ii) Full ambiguity

If the unfolding speech does not appositely resolves NS anaphora, then there is full ambiguity, often left unresolved, allowing more than one possible interpretation of the subject referent (i.e. being ambiguous between TC and TS) or an erroneous interpretation (e.g. TC instead of TS). Two types of full ambiguity were observed in the production data: morphological ambiguity, due to ambiguous morphology, morphological errors (e.g. agreement violations in verbs, clitics and case assignment) or omission errors; and genuine ambiguity, which was triggered by the NS itself.

Morphological ambiguity

Morphological ambiguity, found solely in the bilingual groups (N=10), was triggered by mishandling of morphological phenomena, which were the following (N of cases per group in brackets):
- ambiguous morphology of the verb *ime* 'to be/exist' (IMM: 2, HS: 1, L2: 1)
- morphological errors in verb endings (subject-verb agreement) (HS: 2, L2: 1)
- misuse of gender assignment in clitics (HS: 2)
- omitting an obligatory clitic pronoun (HS: 1)

An example of morphological ambiguity is (85).

(85) *[...] ke o skilos pu zuse se afto to oreo to kipo iðe pu i kaki- i kakia ɣata ithele na fai ta pulakia, ∅? **piðikse** apano **tu**, ∅? **ton** pire ap' tin ura, ∅? **ton** petakse kato.*

'[...] and the dog who lived in this nice garden saw that the bad- the bad cat wanted to eat the little birds, **[s/he?] jumped** on **him**, **[s/he?]** grabbed **him** from the tail, **[s/he?]** threw **him** down.'

<div align="right">(S31, HS, age: 44, near-native)</div>

In (85), the context in *piðikse* ('jumped') is TS, but a TC interpretation is biased: the correct subject referent is the dog but it appears to be the cat. The trigger of this ambiguity is not the NS itself but the wrong gender agreement in the clitic(s). The chosen nouns to refer to the dog (*o skilos*) and the cat (*i kakia ɣata*) are masculine and feminine respectively. Although the listener is initially inclined to assign the dog (*o skilos*) to the NS of *piðikse* ('jumped'), the speaker made morphological errors in clitics using masculine instead of feminine gender. Thus, apparently the cat is the subject referent which jumped on the dog. Due to the (erroneous) gender cues in clitics, this case of ambiguity remains unresolved, leading to miscommunication.

Although proficiency plays a vital role in handling morphology accurately, this is not neatly evidenced in the eight cases of morphological ambiguity produced by the HS/ L2ers. There was at least one case of such ambiguity in each proficiency level. Three out of four occurrences of this type of ambiguity found in the basic proficiency level were produced by one speaker (S30). There were also two cases detected in the intermediate level, one case in the advanced and one case in the near-native levels. Given the low number of tokens and speakers, it cannot be inferred from these data that proficiency is clearly associated with morphological ambiguity (see Table 6.22).

Genuine ambiguity

In genuine ambiguity, the referent of the NS in TS was not resolved because of the presence of more than one possible antecedents and the concurrent absence of relevant morphological, semantic or contextual cues. Contrary to ambiguity caused by morphological issues, genuine ambiguity involved cases of NS in TS which were not properly disambiguated due to subject omission. It was thus impossible or

difficult to establish the correct referent of the NS. There were 18 cases of unresolved NS in TS, with 8 cases found in immigrants produced by 5 speakers, 2 cases in HS by 2 speakers, and 7 cases in L2ers by 6 speakers (examples 86-88). In monolinguals, there was only one case in the Greek data.

(86) *Espase to poði tu, ala i ajelaða lipon ixe ena filo, itane ena peristeri, itane nosokoma. Ki efere lipon to asthenoforo. Irthe ena asthenoforo ki ∅? **efere** lipon ta tiafta, to lefkoplasti, ke tu ∅? **eðese** to podi lipon.*

'[He] broke his leg, but the cow then had a friend, [she] was a dove, [she] was a nurse. And [she] brought then the ambulance. The ambulance came and **[she/it?]** brought then stuff, the bandage, and so **[she/it?]** tied his leg.'

(S51, IMM, age: 76, LoR: 63)

In (86), the two ambiguous NS in the verbs *efere* ('brought') and *eðese* ('bound') appear as having the 'ambulance' (*asthenoforo*) as their subject antecedent following a TC interpretation. However, according to the storyline the subject/agent in *efere* is the 'dove' (*peristeri*) while in *eðese* it is the 'cow' (*ajelaða*). LS would be required for the listeners to interpret correctly the subject antecedents in both cases.

In (87), the NS in *ftiaxni* ('fixes') is ambiguous. The contextual cues given by the speaker indicate that the bird fixes the horse's leg. However, according to the story, the subject referent is the cow. A LS would be necessary to avoid misinterpretation.

(87) *Lipon, piðai to aloyo to fraxti, ala pefti. Ke ta ðio zoakia pu itan eki pane konda tu ke to- to pulaki, eh, ixe ena valitsaki proton voithion ke mazi me tin ajelaða tu-... Fenete espase to poðaraki tu, ekane kapja pliji sto poðaraki tu ke ∅? tu **ftiaxni** to poðaraki.*

'So, the horse jumps the fence but [he] falls. And the two little animals that were there go close to him and the- the little bird, eh, had a first-aid kit and together with the cow-... It seems that [he] broke his leg, [he] had some wound in his leg, and **[s/he?]** fixes his leg.'

(S25, HS, age: 48, near-native)

In (88), the context in *travai* ('pulls') appears to be TC and not TS. Due to the NS, the referent is apparently the 'little bird' (*to pulaki*). However, the picture-story shows that the dog pulls the cat. An overt subject (OSP or LS) would disambiguate the reference.

(88) *[...] ke arxizi o γato- o γata na kitai ta tria pulakia pano sto ðentro ke arxizi na-na aneveni ke ton kitai enas skilos. Ki oso aneveni o- i γata, ti travai tin ura. Ki erxete o- to pulaki, to- i mitera to pulaki, ksero 'γo, me ena skuliki ja fajito ke Ø* tin **travai** ke i γata fevji jati o piso tin kiniγai ti γata ke afta.*

'[...] and the cat- the cat starts looking at the three little birds on the tree and [he] starts climbing and a dog sees him. And as the cat climbs, [he] pulls her tail. And the- the little bird, the mother little bird, comes with a worm for food and **[s/he?]** pulls her and the cat leaves because the one behind her chases the cat and that is all.'

(S16, L2, age: 35, advanced)

Observations

Ambiguity was neither always unresolved (temporary ambiguity) nor always triggered by the NS (see some cases of morphological ambiguity). It was also demonstrated how agreement errors may interfere with establishing coreference. Full ambiguity can be reasonably considered more problematic than temporary ambiguity, but it cannot be established whether such cases ambiguity, either morphological or genuine, were performance errors. Table 6.26 shows the number and percentage of the fully ambiguous clauses involving NS in TS in relation to the total number of NS in TS, as well as the number and percentage of speakers that produced such ambiguities.

Table 6.26. Use of fully ambiguous NS in TS contexts in all groups

Group	Fully Ambiguous		Speakers using ambiguous NS in TS	
	N	%	N	%
Spanish	0/59	*0*	0/20	*0*
Greek	1/60	*1.7*	1/20	*5*
IMM	10/115	*8.7*	8/35	*22.9*
HS	8/67	*11.9*	6/21	*28.6*
L2	9/42	*21.4*	8/20	*40*

The age of speakers was significantly associated with production of ambiguity (§6.3.2.3, Table 6.21), namely, the older the speakers, the more ambiguous NS-TS constructions were produced. Nevertheless, similar cases were not found in older monolingual participants, suggesting that ambiguity is enhanced by bilingualism effects in speaking in the less used/less dominant language.

Proficiency was not clearly associated with production of ambiguous NS in TS in the groups of HS and L2ers (§6.3.2.3, Table 6.22). However, the cases of morphological ambiguity can be attributed to low proficiency or, in the case of more proficient speakers, to performance errors.

In this respect, genuine ambiguity, triggered by actual subject omission, is the most problematic case of ambiguity since it generally remains unresolved, does not directly involve morphological mishandling, which could signal low proficiency and/or performance errors, and concerns particular uses of infelicitous NS. Table 6.27 displays only the cases of genuinely ambiguous NS in TS.

Table 6.27. Use of genuinely ambiguous NS in TS contexts in all groups

Group	Genuinely Ambiguous		Speakers using ambiguous NS in TS	
	N	%	N	%
Spanish	0/59	*0*	0/20	*0*
Greek	1/60	*1.7*	1/20	*5*
IMM	8/115	*7*	5/35	*14.3*
HS	2/67	*3*	2/21	*9.5*
L2	7/42	*16.7*	6/20	*30*

The L2 group manifested the greatest proportion of genuinely ambiguous NS in TS (16.7%) as well as the greatest proportion of speakers producing them (30%). The L2 participants who produced such instances of ambiguity were speakers of basic proficiency (N=3) (S28, S29, S57) or Chileans with advanced/near-native level in Greek who had lived in Greece for long periods of time (N=3) (S15, S16, S17). In the immigrants' group, the age effect was immediately evident since all genuinely ambiguous NS were produced by speakers of 76 years old and above (S09, S32, S51, S56, S146). Only two HS produced such cases of ambiguity (S25, S42).

6.3.2.4 Summary and discussion

In TS contexts, monolinguals and bilinguals expressed change of subject referent by mostly using LS. The HS group produced significantly more NS and significantly fewer LS in TS than both Greek monolinguals and L2ers. NS were used in TS contexts at higher rates than those of OSP. Ambiguity was triggered by the NS per se or by mishandling of other morphological phenomena in the bilingual performance, while ambiguity rates were negligible in monolinguals. OSP were used in TS by all groups of speakers, but were found to be significantly more frequent in Spanish monolinguals than in the Greek-speaking groups. No overproduction of OSP was manifested in bilinguals, whose performance was similar to the Greek monolingual. Specifically, OSP were sparsely produced with no differences between the groups.

In sum, differences in subject rates in TS were found in Spanish vs Greek-speaking groups in OSP expression and in HS vs Greek monolinguals and L2ers in NS and LS expression (§6.3.2.1). OSP seemed to have a wider scope in Spanish than in Greek. Leaving details aside, all groups' performance conformed to accessibility propositions (Ariel 1990). Accordingly, fuller referring expressions corefer to less salient/prominent antecedents. Bilinguals, however, tended to be more ambiguous than monolinguals when employing NS in TS. Although NS in TS are not a priori infelicitous, the relatively high frequency of ambiguous NS in TS by bilingual speakers is a finding which indicates a difference between bilinguals and monolinguals in the expression of reference. These findings are further discussed in §6.6.

6.3.3 Ambiguous verb morphology in Spanish

Focusing on verb forms with ambiguous inflection between first and third person in Spanish (see §2.3.7.2), the production data showed that OSP were used only in TS contexts at a similar rate to that of NS (21.95% vs 19.51% respectively). Ambiguous verbal morphology in TC clauses did not trigger OSP (Figure 6.15, Table 6.28). The present data, therefore, indicate that OSP were not used more often with verbs bearing ambiguous inflection in Spanish.
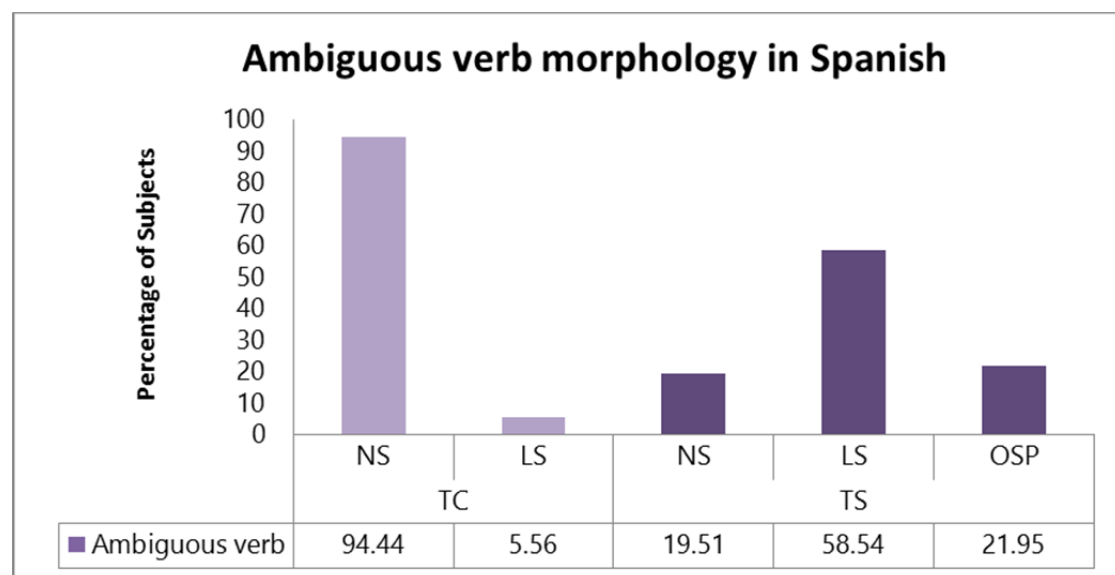
Figure 6.15. Use of subjects with verbs with ambiguous verbal morphology in Spanish

Table 6.28. Use of subjects with verbs with ambiguous verbal morphology in Spanish

Spanish	TC		TS		
	NS	LS	NS	LS	OSP
Ambiguous verb	34	2	8	24	9
morphology	*94.44*	*5.56*	*19.51*	*58.54*	*21.95*

Drawing on the above data, as well as on Filiaci (2011) and Van Esbroeck (2014), it is assumed that ambiguous verb morphology does not affect NS vs OSP distribution in Spanish. This is also relevant for the AR task of the present study, which used verbs in past imperfect in the experimental items (see Chapter 7).

6.4 Multinomial logistic regressions

This section presents the results obtained from multinomial logistic regressions applied in the production data in order to examine the statistical significance of the findings in more detail as well as to explore the associations between linguistic and sociolinguistic variables. The analysis included three independent sociolinguistic variables:

(a) Group of speakers (Spanish monolinguals, Greek monolinguals, IMM, HS, L2)
(b) Age at testing (Age)
(c) Proficiency (only for HS and L2)

The association of these variables with Category of subjects (NS, OSP, LS) as the dependent variable was examined in two contexts: TC and TS.

Four statistical models were developed with the Category of NS used as the baseline against which the other categories of subjects were compared in all models. The analyses examined TC and TS separately corresponding to the four models as follows:

(a) First model: analysis of Group of speakers and Age in association with Category of subjects in the monolingual groups (Greek, Spanish).

(b) Second model: analysis of Group of speakers and Age in association with Category of subjects in the Greek-speaking groups (Greek monolinguals, IMM, HS, L2) with Greek monolinguals being the baseline group.

(c) Third model: analysis of the variable Group of speakers and Age in association with Category of subjects in the three bilingual groups (IMM, HS, L2) with immigrants being the baseline group.

(d) Fourth model: analysis of the variable Group of speakers, Age and Proficiency in association with the Category of subjects in the HS and L2 groups.

Relative Risk (RR), Standard Error of the RR (SE), z-value of the model, p-value of the model and 95% Confidence Interval (CI) of RR are shown in the regression tables.

6.4.1 Topic Continuity

6.4.1.1 First model: Spanish and Greek monolingual groups

A multinomial logistic regression was conducted in order to analyse the variables Group of speakers and Age in the two monolingual groups (Greek, Spanish) in association with the Category of subjects (N of observations: 1044). Spanish monolinguals were the reference group. The results showed a pseudo R^2=0.022 and a significant model (p=0.040).

Age showed a significant (p=0.016) and negative association (RR=0.969) with LS compared to NS. This means that being older by one year decreases by 3.2% the chance of producing LS over NS in TC contexts. A further multinomial logistic regression conducted for each of the two monolingual groups separately showed that Age was significantly associated only in the Greek monolingual group, not in the Spanish group. In particular, there was a significant (p=0.033) and negative association (RR=0.953) between Age and LS over NS, i.e. older Greek monolinguals tended to use less LS than NS in TC. There was no significant association between the two monolingual groups with regard to Category of subject (Table 6.29).

Table 6.29. Multinomial logistic regression: subject category in TC in monolinguals

Category of Subject	RR	SE	z	p	RR CI LB	RR CI UB
OSP						
Age	1.031	0.039	0.820	n.s.	0.958	1.111
Group of speakers						
Greek	0.412	0.506	-0.720	n.s.	0.037	4.573
LS						
Age	0.969	0.013	-2.410	0.016	0.944	0.994
Group of speakers						
Greek	0.690	0.200	-1.280	n.s.	0.391	1.217

(RR=Relative Risk, SE=Standard Error, RR CI=Relative Risk Confidence Interval, LB=Lower Bound, UB=Upper Bound)

The results showed no differences in the use of subjects between the monolingual groups in TC contexts. It was found that Age plays a role in subject expression in the Greek monolinguals: the older the participants the more frequent the use of NS in TC.

6.4.1.2 Second model: Greek-speaking groups

A multinomial logistic regression was conducted in order to analyse the variables Group of speakers and Age in the Greek-speaking groups (Greek monolinguals, IMM, HS, L2) with Greek monolinguals being the reference group (N of observations: 2161). The results showed a pseudo $R^2=0.038$ and a significant model (p<0.001).

Age showed a significant (p=0.038) and negative association (RR=0.989) with LS compared to NS, i.e. being older by one year decreases by 1% the chance of using LS over NS in TC contexts. A further multinomial logistic regression showed that the association between Age and LS was significant only in the case of L2 (p=0.004, RR=0.970) and Greek monolinguals as shown in the previous section. The groups of HS and L2ers showed a significant (p=0.015, p<0.001) and positive association (RR=2.050, RR=3.342) with LS compared to Greek monolinguals. Specifically, the HS and the L2 group used 2.1 times and 3.3 times more LS than NS in TC respectively. Age and Group of speakers showed no significant association with OSP (Table 6.30).

Table 6.30. Multinomial logistic regression: subject category in TC in Greek speakers

Category of Subject	RR	SE	z	p	RR CI LB	UB
OSP						
Age	0.984	0.020	-0.80	n.s.	0.945	1.024
Group of Speakers						
IMM	1.381	1.693	0.26	n.s.	0.125	15.265
HS	1.863	2.644	0.44	n.s.	0.115	30.095
L2	8.270	9.083	1.92	n.s.	0.959	71.222
LS						
Age	0.989	0.005	-2.08	0.038	0.979	0.999
Group of Speakers						
IMM	1.163	0.313	0.56	n.s.	0.687	1.971
HS	2.050	0.608	2.42	0.015	1.147	3.665
L2	3.342	0.886	4.55	<0.001	1.987	5.620

(RR=Relative Risk, SE=Standard Error, RR CI=Relative Risk Confidence Interval, LB=Lower Bound, UB=Upper Bound)

The immigrants did not differ from the Greek monolinguals in TC. The HS and L2ers tended to use more LS in TC than Greek monolinguals, as also seen in §6.3.1. Moreover, the use of NS increases with age in L2ers, similarly to monolinguals.

6.4.1.3 Third model: Bilingual groups

A multinomial logistic regression was conducted in order to analyse the variables Group of speakers and Age in the three bilingual groups (IMM, HS, L2) with immigrants being the reference group (N of observations: 1591). The results showed a pseudo R^2=0.036 and a significant model (p<0.001).

There was a significant and positive association of LS with the HS group (p=0.033, RR=1.750): the HS produced 1.7 times more LS than NS compared to immigrants. Moreover, a significant and positive association of both OSP (p=0.34, RR=5.934) and LS (p<0.001, RR=2.910) was found with the L2 group. The L2ers produced 6 times more OSP and 3 times more LS than NS compared to immigrants in TC. Age did not reveal significant associations with respect to Category of subject (Table 6.31).

Table 6.31. Multinomial logistic regression: subject category in TC in bilinguals

Category of Subject	RR	SE	z	p	RR CI	
					LB	UB
OSP						
Age	0.980	0.021	-0.90	n.s	0.939	1.023
Group of speakers						
HS	1.370	1.690	0.26	n.s.	0.122	15.300
L2	5.934	4.991	2.12	0.034	1.141	30.850
LS						
Age	0.992	0.005	-1.49	n.s.	0.981	1.002
Group of speakers						
HS	1.750	0.532	2.13	0.033	1.050	2.921
L2	2.910	0.658	4.66	<0.001	1.860	4.550

(RR=Relative Risk, SE=Standard Error, RR CI=Relative Risk Confidence Interval, LB=Lower Bound, UB=Upper Bound)

The HS and L2ers were distinguished from the immigrants in producing more LS in TC. The L2 produced more OSP in TC than the immigrants; however, the number of OSP in the L2 group was low (N=5) and mostly produced by one speaker.

6.4.1.4 Fourth model: HS and L2ers

A multinomial logistic regression was conducted in order to analyse the variable Group of speakers, Age and Proficiency in the groups of HS and L2ers (N of observations: 721) with HS as the reference group. The results showed a pseudo $R^2=0.063$ and a significant model (p<0.001).

A significant and negative association (p=0.008, RR=0.131) between Proficiency and OSP was found when considering both groups. Specifically, being of higher proficiency by one level decreases the chance of using OSP over NS by 7.6 times in TC. Age showed a significant and negative association (p=0.037, RR=0.920) with OSP relative to NS, suggesting that being older by one year decreases by 8% the chance of using OSP over NS in TC. The results indicate that being older and more proficient in Greek conduces to less frequent use of OSP in TC. Age also showed a significant and negative association (p=0.001, RR=0.973) with LS compared to NS: being older by one year decreases by 2.7% the chance of using LS over NS in TC. A further regression showed that this association emerged in the L2 group (p=0.004, RR=0.970) and not in HS. Proficiency was not associated with LS (Table 6.32).

Table 6.32. Multinomial logistic regression: subject category in TC in HS and L2

Category of Subject	RR	SE	z	p	RR CI	
					LB	UB
OSP						
Age	0.920	0.038	-1.83	0.037	0.850	0.994
Proficiency	0.131	0.100	-2.82	0.008	0.030	0.590
Group of speakers						
L2	1.360	1.612	1.11	n.s.	0.133	13.896
LS						
Age	0.973	0.008	-3.01	0.001	0.950	0.990
Proficiency	1.110	0.140	0.47	n.s.	0.864	1.423
Group of speakers						
L2	1.500	0.400	1.40	n.s.	0.890	2.533

(RR=Relative Risk, SE=Standard Error, RR CI=Relative Risk Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Frequency of NS in TC increased along with age in this context. In addition, the more proficient HS and L2ers tended to use OSP less often than the less proficient speakers, although this finding is based on insufficient data to be generalised.

6.4.2 Topic Shift

6.4.2.1 First model: Spanish and Greek monolingual groups

A multinomial logistic regression was conducted in order to analyse the variables Group of speakers and Age in the monolingual groups (Greek, Spanish) in association with the Category of subjects (N of observations: 490). Spanish was the baseline group. The results showed a pseudo R^2=0.018 and a significant model (p=0.013).

A significant and negative association (p=0.012, RR=0.259) was found in OSP compared to NS between the two monolingual groups. This means that being Greek monolingual decreases the chance of use of OSP over NS in TS contexts by 3.9 times by comparison with Spanish monolinguals. Age was not significantly associated with the use of OSP. Additionally, the variables Age and Group of speakers did not show significant association with LS (Table 6.33).

Table 6.33. Multinomial logistic regression: subject category in TS in monolinguals

Category of Subject	RR	SE	z	p	RR CI LB	UB
OSP						
Age	1.009	0.017	0.540	n.s.	0.977	1.043
Group of speakers						
Greek	0.259	0.139	-2.520	0.012	0.091	0.739
LS						
Age	1.009	0.008	1.050	n.s.	0.992	1.025
Group of speakers						
Greek	1.219	0.260	0.930	n.s.	0.803	1.851

(RR=Relative Risk, SE=Standard Error, RR CI=Relative Risk Confidence Interval, LB=Lower Bound, UB=Upper Bound)

The main finding of this analysis was the difference between the Spanish and Greek monolingual groups with respect to the use of OSP in TS. The former produced OSP in TS more often than the latter, as also shown in §6.3.2.

6.4.2.2 Second model: Greek-speaking groups

A multinomial logistic regression was conducted in order to analyse the variables Group of speakers and Age only in the Greek-speaking groups (Greek monolinguals, IMM, HS, L2) with Greek monolinguals being the baseline group (N of observations: 1043). The results showed a pseudo R^2=0.018 and a significant model (p=0.041).

The group of HS showed a significant and negative association (p=0.011, RR=0.583) with LS in TS compared to Greek monolinguals. This means that being HS decreases by 71.5% the chance of using LS over NS in TS contexts compared to Greek monolinguals (Table 6.34).

Table 6.34. Multinomial logistic regression: subject category in TS in Greek speakers

Category of Subject	RR	SE	z	p	RR CI LB	UB
OSP						
Age	0.994	0.015	-0.400	n.s.	0.964	1.024
Group of speakers						
IMM	0.432	0.299	-1.210	n.s.	0.111	1.680
HS	0.549	0.414	-0.790	n.s.	0.125	2.407
L2	1.143	0.801	0.190	n.s.	0.290	4.511
LS						
Age	0.993	0.004	-1.760	n.s.	0.986	1.001
Group of speakers						
IMM	0.781	0.146	-1.320	n.s.	0.542	1.126
HS	0.583	0.124	-2.530	0.011	0.385	0.885
L2	1.131	0.259	0.540	n.s.	0.722	1.771

(RR=Relative Risk, SE=Standard Error, RR CI=Relative Risk Confidence Interval, LB=Lower Bound, UB=Upper Bound)

The results revealed an association between the HS and LS, indicating that HS are different from Greek monolinguals in using more NS in TS, as also shown in §6.3.2.1.

6.4.2.3 Third model: Bilingual groups

A multinomial logistic regression was conducted in order to analyse the variables Group of speakers and Age in the three bilingual groups (IMM, HS, L2) with immigrants being the baseline group (N of observations: 786). The results showed a pseudo R^2=0.014 and a significant model (p=0.020).

The variable Age revealed a significant (p=0.028) and negative association (RR=0.99) with LS considering the three bilingual groups. This means that being older by one year decreases the chance of using LS over NS by 1% in TS. A further logistic regression revealed that this association was pertinent only in the immigrants' group (p=0.019, RR=0.99) and not in the HS and L2 groups.

There were no significant associations regarding OSP. In addition, the HS and L2 group were not significantly associated with Category of subjects in TS (Table 6.35).

Table 6.35. Multinomial logistic regression: subject category in TS in bilinguals

Category of Subject	RR	SE	z	p	RR CI LB	UB
OSP						
Age	0.987	0.016	-0.71	n.s.	0.955	1.021
Group of speakers						
HS	1.263	0.984	0.30	n.s.	0.274	5.821
L2	2.570	1.885	1.29	n.s.	0.610	10.827
LS						
Age	0.990	0.004	-2.19	0.028	0.982	0.999
Group of speakers						
HS	0.742	0.140	-1.57	n.s.	0.513	1.076
L2	1.429	0.298	1.71	n.s.	0.950	2.151

(RR=Relative Risk, SE=Standard Error, RR CI=Relative Risk Confidence Interval, LB=Lower Bound, UB=Upper Bound)

The variable Age was significantly associated with the use of LS. The older bilingual speakers were more likely to use NS in TS contexts compared to the younger speakers and this was triggered by the performance of the immigrants.

6.4.2.4 Fourth model: HS and L2ers

A multinomial logistic regression was conducted in order to analyse the variables Group of speakers, Age and Proficiency in the groups of HS and L2ers (N of observations: 390). The results showed a pseudo R^2=0.046 and a significant model (p<0.001).

A significant (p=0.005) and negative association (RR=0.150) of Proficiency with OSP was found considering both groups. This means that the higher the proficiency of the speakers the lower the rate of OSP use over NS in TS contexts: being more advanced by one level decreases by 6.7 times the chance of using OSP over NS in TS. However, the number of OSP was very low considering only the groups of HS and L2ers. Additionally, there was a significant positive association (p=0.003, RR=2.013) between the L2 group and LS, indicating that being an L2 speaker increases the chance of using LS over NS by 2 times compared to HS. Age was not significantly associated with Category of subjects. Group of speakers was not significantly associated with OSP. Proficiency did not show a significant association with LS in TS (Table 6.36).

Table 6.36. Multinomial logistic regression: subject category in TS in HS and L2

Category of Subject	RR	SE	z	p	RR CI LB	UB
OSP						
Age	0.946	0.030	-1.72	n.s.	0.887	1.007
Proficiency	0.150	0.101	-2.81	0.005	0.399	0.563
Group of speakers						
L2	0.835	0.716	-0.21	n.s.	0.156	4.478
LS						
Age	0.998	0.008	-0.33	n.s.	0.983	1.013
Proficiency	1.069	0.125	0.58	n.s.	0.851	1.344
Group of speakers						
L2	2.013	0.479	2.94	0.003	1.264	3.208

(RR=Relative Risk, SE=Standard Error, RR CI=Relative Risk Confidence Interval, LB=Lower Bound, UB=Upper Bound)

The L2ers produced more LS in TS than the HS. Additionally, as also found in the TC contexts, OSP were used less frequently by the more proficient speakers, but this finding is based on insufficient data.

6.4.3 Age

In order to explore the effect of Age over Category of subjects in the oral performance of the groups of speakers, further multinomial logistic regressions were performed introducing Age as independent variable in each group of speakers separately. Tables 6.36 and 6.37 display the results of these regressions for TC and TS only respecting LS, because LS were found to be significantly associated with Age.

Age showed a significant and negative association with LS in the group of monolingual Greek (p=0.033, RR=0.953) and L2ers (p=0.004, RR=0.969) in TC. This means that being older decreases the chance of using LS over NS by 4.9% and 3.1% respectively in TC (Table 6.37).

Table 6.37. Logistic regressions: Age and Category of subject (LS) in TC

	RR	SE	z	p	RR CI	
Spanish	0.978	0.015	0.73	n.s.	0.947	1.124
Greek	0.953	0.211	-2.13	0.033	0.913	0.996
IMM	1.006	0.007	0.98	n.s.	0.993	1.021
HS	0.980	0.013	-1.48	n.s.	0.954	1.006
L2	0.969	0.010	-2.88	0.004	0.948	0.990

In addition, Age showed a significant and negative association in TS with the group of immigrants (p=0.019, RR=0.988), indicating that being older decreases the chance of using LS over NS by 1.2% (Table 6.38).

Table 6.38. Logistic regressions: Age and Category of subject (LS) in TS

	RR	SE	z	p	RR CI	
Spanish	1.006	0.011	0.58	n.s.	0.984	1.029
Greek	1.011	0.012	0.93	n.s.	0.987	1.036
IMM	0.988	0.004	-2.35	0.019	0.978	0.998
HS	1.010	0.010	1.00	n.s.	0.989	1.032
L2	0.985	0.009	-1.52	n.s.	0.966	1.004

The age effect regarding LS in TS in the immigrants' group, which presented substantial variation in age (22-86 years), was examined in a follow-up analysis dividing the group into younger and older speakers, as described in §5.2.5.2.

Pearson Chi-square tests showed a statistically significant difference in LS rates in TS between Younger (N=192, 73.75%) and Older immigrants (N=85, 62.5%) [Pearson χ^2 (1, N=396)=5.35, p=0.021]. Statistically significant difference was also found in their rates of NS in TS [Pearson χ^2 (1, N=396)=4.81, p=0.028], with Older immigrants producing significantly more NS (N=49, 36.03%) than Younger immigrants (N=66, 25.5%). The OSP rate was negligible (N=2 in each subgroup).

In the rates of NS in TS, Greek monolinguals patterned with Younger immigrants (23.35% vs 25.5% respectively). Older immigrants patterned with HS (36.03% vs 34.9% respectively). A statistically significant difference was found between Older immigrants and Greek monolinguals [Pearson χ^2 (1, N=393)=7.13, p=0.008].

6.4.4 Findings

Summarising the findings, Group of speakers revealed significant associations in cases shown in Tables 6.39 and 6.40 and Age in cases shown in Tables 6.41 and 6.42.

Table 6.39. Group of speakers: Summary of significant associations in TC

Topic Continuity		
Association between variables		
Group	**Subject**	**Finding**
HS	LS	The chance of using LS over NS is high compared to Greek monolinguals and IMM.
L2	LS	
L2	OSP	The chance of using OSP over NS is high compared to IMM (biased by the performance of one speaker).

Table 6.40. Group of speakers: Summary of significant associations in TS

Topic Shift		
Association between variables		
Group	**Subjects**	**Finding**
Greek	OSP	The chance of using OSP over NS is low compared to Spanish monolinguals.
HS	LS	The chance of using LS over NS is low compared to Greek monolinguals.
L2	LS	The chance of using LS over NS is high compared to HS.

Table 6.41. Age at testing: Summary of significant associations in TC

Topic Continuity		
Association between variables		
Group	**Subject**	**Finding**
Greek	LS	Being older decreases chance of using LS over NS.
L2	LS	
HS & L2	OSP	Being older decreases chance of using OSP over NS.

Table 6.42. Age at testing: Summary of significant associations in TS

Topic Shift		
Association between variables		
Group	**Subject**	**Finding**
IMM	LS	Being older decreases chance of using LS over NS.

Proficiency revealed significant associations only with respect to OSP in both TC and TS (Table 6.43). As already mentioned, the frequency of OSP occurrences was very low. This means that although there was significant association between OSP use and proficiency, these results are not conclusive.

Table 6.43. Proficiency: Summary of significant associations in TC and TS

Topic Continuity & Topic Shift		
Association between variables		
Group	**Subjects**	**Finding**
HS & L2	OSP	Being of a higher proficiency decreases the chance of using OSP over NS.

6.5 Summary of findings

With regard to production, the following sections provide the conclusions that can be drawn from the quantitative and qualitative analysis of the narratives corpus.

6.5.1 Greek and Spanish monolinguals

TC contexts were predominantly expressed with NS in both Greek and Spanish without significant differences between the languages. The rates of LS were also comparable. The percentages of NS in TC in matrix clauses were significantly different, with Greek monolinguals producing more NS than Spanish speakers. Their performance was similar in embedded clauses. Qualitatively, both Greek and Spanish monolinguals behaved similarly in TC. The OSP use was both quantitatively and qualitatively similar in the two monolingual groups in TC, i.e. OSP were barely used and their discourse function was to encode focus (emphasis/contrast). The monolinguals' performance in TC was generally along the lines of accessibility theory (Ariel 1990).

TS contexts were largely expressed with LS in Greek and Spanish and the rates of NS were also similar in both languages. Spanish monolinguals used significantly more OSP and significantly fewer LS than Greek monolinguals in this context. Qualitatively, the use of OSP was comparable in the two monolingual groups, i.e. OSP were used to mark TS or focus. Spanish monolinguals expressed TS mostly with LS, but also with NS and to a smaller extent with OSP. Their Greek counterparts expressed TS predominantly with LS, but also with NS and a very low rate of OSP, significantly lower than the OSP rate in Spanish. It was observed that OSP in Greek were used only when they were absolutely necessary for reference disambiguation. The monolingual performance was generally in line with accessibility accounts (Ariel 1990).

Age was significant only in TC and only in the Greek group, not in the Spanish group, with the chance of using NS increasing along with age. The older the monolingual speakers, the wider was the use of NS in TC. This indicates that in general the younger monolingual speakers of Greek tended to be more specific (or over-specific) than the older speakers in expressing TC.

6.5.2 Greek-Spanish bilinguals

In TC contexts, the use of NS was prevalent in the three bilingual groups similarly to monolinguals, but it was considerably higher in Greek monolinguals and immigrants compared to HS and L2ers. The latter groups produced LS in TC significantly more often than the former groups. The qualitative analysis of LS in TC, however, indicated no considerable differences with respect to LS use and redundancy between the groups. There were, therefore, quantitative differences between the groups of Greek monolinguals and immigrants on the one side and HS and L2ers on the other regarding NS and LS use in TC, but overall no qualitative differences were observed in this context. Age was significantly associated with LS the L2 group, with older speakers being prone to use NS over LS in TC more frequently than younger speakers (similarly to Greek monolinguals).

In TS contexts, LS were mostly used by the bilinguals similarly to monolinguals. Age was significantly associated with the use of LS in the immigrants' group: being older increases the chance of using NS over LS in TS. The L2ers produced significantly more LS than the HS in TS. NS were also used in TS by the three bilingual groups, with HS producing significantly more NS than Greek monolinguals and L2ers. The use of OSP in TS by bilinguals was quantitatively and qualitatively similar to the Greek monolingual performance. In contexts of NS in TS, bilinguals produced instances of ambiguity whereas those were negligible in monolinguals. Ambiguity in TS contexts was significantly associated with age: the likelihood that the ambiguous clauses were produced by older participants was higher.

Proficiency in HS and L2ers was associated with OSP use in both TC and TS contexts. Higher proficiency decreases the frequency of OSP. This finding was based on 15 instances of OSP produced by HS and L2ers in all contexts and 5 of these instances were produced by one L2 speaker. Thus, this association is based on evidence which is to some extent biased. Quantitatively, the main findings are (a) overuse of NS in TS by the HS and NS in TS being associated with age in immigrants; and (b) the absence of difference in OSP use in bilinguals groups compared to Greek monolinguals. The findings are discussed more thoroughly in the following section.

6.6 Discussion

The present section discusses the research findings engaging with relevant literature. It is divided into two main sections. Section 6.6.1 discusses the findings obtained from Greek and Spanish monolinguals. Section 6.6.2 discusses the findings obtained from the bilingual groups.

6.6.1 Comparison between Greek and Spanish

The first research question sought to answer whether Greek is different from Spanish in the distribution of third-person subjects in non-focused contexts. The hypothesis was that the two languages would differ in the distribution of OSP. The Spanish OSP was predicted to be more variable than the Greek OSP because of the deictic nature of the latter, which renders it less flexible and less ambiguous by comparison.

The overall pattern of subject distribution in narratives was similar in Greek and Spanish. Considering all contexts, NS were used in more than half the time, LS hovered around 33%[15], while OSP were used sparingly in both languages. This is in line with related research on Greek and Spanish (Dimitriadis 1996; Montrul & Rodríguez Louro 2006; Peskova 2013; Georgopoulos 2017). Both groups favoured the use of NS in embedded clauses obeying general principles of binding: NS of embedded clauses usually maintain the discourse referent introduced in antecedent clauses, hence embedded clauses generally favour NS as the most natural option (Margaza & Bel 2006; Prada Pérez 2009).

Almost half of the embedded clauses were adverbial, this type of embedded being the most commonly used in both groups. The main difference in embedded clauses was the expected use of infinitive in Spanish and the lack thereof in Greek, resulting in subjunctive complement clauses being much more frequent in Greek than in Spanish. Subject distribution, however, was not found to be affected by the difference in complement clauses with infinitive and subjunctive verb forms (see §6.2.2).

---

[15] Percentages are rounded in the current section for convenience purposes.

Preliminary overall differences involved the use of NS, which was significantly high in Greek compared to Spanish (61.7% vs 56.5%), and the use of OSP, which was significantly high in Spanish compared to Greek (2.7% vs 0.6%). The rates of LS were similar in Greek (33%) and Spanish (32.6%) considering all contexts. This was an initial indication that the difference in the two languages might lie in the frequency of NS and OSP, which were further examined within the particular contexts of use in order to assess their scope.

6.6.1.1 Topic continuity

When the same subject was maintained in the discourse, the context involved TC and was predominantly expressed with NS in both Greek and Spanish (96% and 94% respectively). This is the default mapping in line with previous studies, such as Blackwell (2003: 246), who affirms that in narratives, due to antecedent saliency, 'the speakers produce the most minimal form available to them to express coreference'. The groups of Greek and Spanish monolingual speakers were not considerably different in the use of LS and NS in TC contexts on the whole. Only in matrix clauses of TC, Greek monolinguals produced significantly more NS (95%) and fewer LS (5%) than Spanish monolinguals (87% and 12% respectively). This may imply that NS are used more widely in Greek in TC than in Spanish or that subject expression in TC in Spanish is more explicit than in Greek in matrix clauses.

Lexical subjects

LS were produced in TC by both monolingual groups to similar degrees (Greek: 4%, Spanish: 6%). LS in subject maintenance are considered to be overexplicit, hence redundant, because the antecedent is already salient. Qualitative analyses showed that around 78% of LS in TC in Greek and Spanish were not redundant and only approximately 22% of LS in TC were actually superfluous in each group (see §6.3.1.2). This suggests that TC is not strictly encoded with NS. Such findings raise the question of what induces use of LS in TC in monolinguals, given that utterances should not convey unnecessary information and LS are unnecessary in contexts of subject maintenance.

Qualitative analyses of the TC contexts shed light on how monolingual speakers use LS as a stylistic means to restate the subject-topic of the utterance without causing redundancy. It was illustrated in §6.3.1.2 that non-redundant hence felicitous LS were used in TC following stylistic conventions and norms of narrating. Blackwell and Lubbers Quesada (2012) assert that there are additional pragmatic factors which constrain or trigger the use of non-null forms. Lexically full or explicit referring expressions can be used 'to signal the beginning of a new scene or action, reiteration of a current topic, contrastive focus, or another type of emphasis' (Blackwell & Lubbers Quesada 2012: 154; see also Blackwell 2003). Despite TC, in cases of non-redundant LS, 'the narrative structure appears to require a type of emphasis to carry the story-line forward' (ibid: 155). In the present study, most cases of LS in TC in both monolingual groups were produced succeeding existential-presentational clauses (e.g. *Once upon a time there was a horse. The horse*...). In other words, LS were used anaphorically immediately after the first mention of the referent. According to Blackwell (2003: 229), 'this marked expression may be explained, at least in part, by the thematic prominence of the actor in the local discourse'. In Spanish, in particular, the proportion of this type of non-redundant LS was found in 61% of cases of LS in TC (N=17). In Greek, it was found in 39% of such cases (N=9). This difference seems to reflect a narration pattern which was more prevalent in Spanish than in Greek and could also partly explain the previously-mentioned quantitative difference in the proportions of LS and NS in matrix clauses between the groups.

Redundancy

As regards the redundant uses of LS in TC, the phenomenon can be explained considering three potential factors. First, redundancy may be (partly) due to the instructions received by the participants, according to which they should narrate the stories to an imaginary listener who was ignorant of the plots, therefore they should be explicit and unambiguous. Language users generally determine the optimal output from both their own perspective and the perspective of their addressee. Responsiveness to the instructions of the task thus may have given rise to some over-explicitness in subject expression taking into account the hypothetical listener's

perspective. In other words, LS in TC may be related to listener-oriented referential choices (Hendriks et al. 2014). In addition, monolingual native grammars manifest some tolerance for redundancy, as reported in several studies (e.g. Sorace et al. 2009; Georgopoulos 2017; Lozano 2018). Hence, LS may be redundantly used to some extent by monolingual speakers, especially in informal and spontaneous oral speech uttered without prior (careful) planning.

Redundancy may be also explained in terms of degrees of referent salience or accessibility (Givón 1983; Ariel 1990; Arnold et al. 2013). Arnold and Griffin (2007) showed that speakers are more likely to use explicit referring expressions when there is more than one character present in the discourse context. The presence of additional characters makes the topic-character less prominent/accessible since it competes for attention in the speaker's discourse representation. Given that reference production is conditioned by attentional resources, 'when two characters are present in the discourse, they share the attentional resources available and each receives less activation in the speaker's internal representation. As a result, speakers are less likely to use an attenuated expression [...] to refer to that character' (Arnold & Griffin 2007: 528). Most of the scenes of the narratives in the present study involved at least two characters taking part in the events. Speakers were also assumed to be sensitive to presence of additional characters even if they were to appear subsequently. Accessibility accounts may thus explain to some extent the redundant use of LS in TC contexts in monolinguals, which in any case was low.

Overt subject pronouns

When OSP do not produce TS, they are commonly associated with the discourse function of focus. OSP in TC were scantly used by monolinguals, with only two occurrences in Spanish and one in Greek. In these cases, OSP encoded contrast or emphasis, thereby conveying focused information. This is in accordance with the literature: 'in contexts where co-reference is maintained, the use of the full pronoun in a sense "warns" the addressee that another kind of contrast is intended by the speaker. In these contexts, the use of *aftos* promotes an emphatic and contrastive, contrary-to-expectation reading' (Chiou 2012: 51; see also Tsimpli 2011; Prentza &

Tsimpli 2012; Lozano 2018). It was anticipated that in Spanish there might be some redundant OSP in TC, since some tolerance for redundancy thereof has been reported (e.g. Alonso-Ovalle et al. 2002; Keating et al. 2011; Georgopoulos 2017; Lozano 2018). However, this was not the case in the Spanish production in this study.

The age factor

A significant association revealed that the older the Greek monolinguals the fewer LS were used in TC, i.e. they used more NS than younger speakers. This finding implies that the younger participants were more explicit in expressing subject maintenance compared to older speakers. The observed behaviour may be ascribed to the aforementioned factors which plausibly influence production of LS in TC, namely pragmatic factors, responsiveness to task instructions, tolerance for redundancy and accessibility effects. Similar differences between younger and older speakers, with the former being more informative than the latter in TC, are also reported in Hendriks et al. (2014). Younger monolinguals seem to be extra sensitive to interpretative considerations. The older Greek monolingual speakers, showing a preference for NS in TC, chose the most economical option (the NS) more often, which, albeit felicitous, implies reduced monitoring in narratives production compared to younger speakers. Such association between subject category and age was not observed in Spanish.

Conclusion

Overall, the expression of TC was according to predictions, with no substantial differences between Greek and Spanish. NS were mainly used for subject maintenance. The preference for NS in TC serves to avoid redundancy in discourse (Gelormini-Lezama & Almor 2011; Shin & Cairns 2012). The rate of overt subjects was low (LS) or insignificant (OSP). Thus, in both languages there was a tendency for producing NS over overtly realised subjects when coreference was intended. Subject distribution in TC generally obeyed the discourse-pragmatics constraints which regulate null and overt options in NS languages. The findings were along the lines of accessibility accounts (e.g. Ariel 1990) and several other studies (e.g. Montrul & Rodríguez Louro 2006; Miltsakaki 2007; Mayol 2012; Montrul 2016a).

6.6.1.2 Topic shift

Contexts involving a shift in subject/referent from the subject/referent of the previous clause are contexts of TS. Monolingual Greek and Spanish speakers encoded TS by predominantly using LS as often as 75% and 67% respectively. This is generally in line with related literature (see Lozano 2009). The difference in the aforementioned relative frequencies was statistically significant. Both groups also used a considerable amount of NS in TS, which was quantitatively similar: 23% in Greek and 25% in Spanish. A significant difference arose in the rates of OSP which was lower in Greek (2%) and higher in Spanish (8.2%). Since the rates of NS in TS were similar while the rates of LS and OSP were statistically different, the crosslinguistic difference seems to lie in the production of overt forms in TS. While Spanish overtly encodes TS with LS and also OSP to a lesser degree, Greek does so by using mostly LS and significantly fewer OSP than Spanish. Greek monolingual speakers seem to 'counterbalance' the restricted use of OSP in TS by instead using LS, whose production is more frequent in Greek than in Spanish in this context. Aside from differences in relative frequencies, OSP were used in both languages to mark TS sometimes combined with focus.

Null subjects

NS were noticeably used in TS contexts 23% and 25% of the time in Greek and Spanish, similarly to other studies (e.g. Blackwell 2003; Pinto 2014), without causing ambiguity. The rates of NS in TS in the monolingual groups were similar to those attested in Pinto (2014) for the same context in monolingual Italian (20%). The anaphora pattern of a NS as a minimal expression in non-coreference was possible by mutual knowledge allowed by inferences as well as by cues in the context permitting grammatical and/or pragmatic recoverability (see Ariel 1990). The omission of overt subjects in TS was thus guided by structural and contextual conditions without resulting in infelicity.

Overt subject pronouns

The relative frequency of OSP was low in the monolingual groups in all contexts (Greek: 0.6%, Spanish: 2.7%) and in TS in particular (Greek: 2%, Spanish: 8.2%). This is line with related literature (e.g. Dimitriadis 1996; Shin & Cairns 2012), but raises the question of why OSP are so infrequent in the performance of monolinguals (and bilinguals). Some factors are considered below.

(a) As seen in Chapter 2, the paradigms of verb inflection in both languages are rich, i.e. every number/person combination has a different suffix; consequently, the third person is distinguished uniquely (with some exceptions of inflectional syncretism). It follows that if recoverability is not at stake, OSP can be left out. This complies with the Avoid Pronoun Principle, which states that OSP should be avoided unless their realisation has some semantic or pragmatic effect.

(b) According to Peskova (2013), the low frequency of third-person pronouns is attributed to their anaphoric nature, i.e. to their interpretation as given information, and to the fact they are used in narrative (as opposed to interactive speech). Pinto (2014) argues that contextual and grammatical cues (beyond verbal inflection) often render OSP pleonastic in NS languages. Arnold et al. (2009) also deem that full LS are easier to produce than pronouns, since the latter must be licensed by the context. These accounts related to context and the referents' salience/accessibility may explain the attested low use of OSP in TS.

(c) As mentioned in §6.3, within the total of subject referents considered in the study, 1% of subjects involved inanimate entities, which are more often referenced by LS or NS than with OSP (e.g. Luján 1999b; Lozano 2009). In addition, 6.6% of subjects were in plural number, which has been found to trigger OSP less often than singular (e.g. Lozano 2009; Prada Pérez 2018). In addition, OSP are rarely used in embedded TC contexts when the embedded subject is coreferential with the matrix clause; hence, in some related studies such embedded contexts are excluded (Prada Pérez 2018). Including inanimate and plural referents as well as TC embedded clauses may have affected the relative frequency of OSP compared to the other subject forms.

(d) With respect to Greek, some of the plausible causes of the relatively 'rare' occurrence of third-person singular OSP were explored in §2.3.6.1. The inherent deictic nature of the third-person pronoun *aftos* adds to its markedness resulting in speakers avoiding its use unless there is a particular purpose which justifies it (Dimitriadis 1996; Chiou 2012). Moreover, omission of third-person OSP referring to a person is often regarded as a form of politeness (this being only indirectly related to the particular contexts of the present study).

The last-mentioned properties of the Greek OSP contribute to its infrequency relative to Spanish. On the other hand, Spanish verb inflection exhibits sycretism in third and first person singular in certain paradigms (§2.3.7.2). Shin (2014: 314) argues that 'the tendency to express *él/ella* is related to contextual ambiguity', which is more likely in TS contexts, 'and the potential for such ambiguity is greater with imperfect verbs'. However, the findings in Filiaci (2011) on monolingual Spanish showed that verbal syncretism does not affect pronoun interpretation. Although influence of inflectional ambiguity in OSP expression in Spanish is not clearly established, it could reasonably be a potential factor in production of OSP when alternation of first and third person singular is involved. This, however, is not the case in the present production data. Moreover, the Spanish narratives did not provide evidence suggesting that OSP were produced to compensate for inflectional ambiguity on the verb.

Irrespective of ambiguous verbal morphology, the qualitative analysis showed that even in TS contexts involving competing referents, most of OSP in Spanish (thirteen out of nineteen) would not cause ambiguity if omitted (see also Blackwell 2003). In Greek, on the other hand, all OSP in TS in the spoken corpus (N=5) were clearly needed to remove referential ambiguity. Thus, the deictically marked OSP in Greek is used with restraint, i.e. only when absolutely necessary, while the Spanish personal pronoun does not bear an equally strong deictic component and is thus somewhat more variable, i.e. more liberally used in TS. These facts may explain the difference in the relative frequency of OSP in TS between Greek and Spanish. It was therefore observed that differences in OSP between the two languages as attested in narrative production of monolinguals seem to reside in the way OSP are used at the discourse pragmatic level only in TS and not in TC contexts.

Conclusion

Overall, the expression of TS was according to predictions, with differences between Greek and Spanish found in the relative frequency of OSP, being higher in Spanish and lower in Greek. This may relate to the fact that the Greek OSP is deictically marked, while the Spanish OSP is deictically less marked. LS were mainly used in TS, followed in frequency by NS and only a low rate of OSP. Subject distribution in TS relied on the discourse-pragmatics factors operating in the two languages. The use of NS was pragmatically appropriate since virtually no ambiguity emerged in TS. It can be thus observed that in both Greek and Spanish there is a tendency for producing LS when non-coreference is intended and, additionally, NS are also felicitously used in this context. The findings are generally in accordance with accessibility accounts (e.g. Gundel et al. 1993; Ariel 1990) as well as with other studies (e.g. Dimitriadis 1996; Montrul & Rodríguez Louro 2006; Miltsakaki 2007; Mayol 2012; Lozano 2018). The deictic interpretation of the Greek OSP seems to be responsible for its more stringent use in narratives, which was evidenced in AR, as will be shown later.

6.6.2 Comparisons of Greek-Spanish bilinguals

The second research question sought to answer whether Greek in contact with Spanish differed from monolingual Greek in the distribution of third-person null and overt subjects in particular contexts; if so, in which contexts and why. Three different types of bilinguals allowed a comparison between monolinguals and bilinguals as well as between bilingual groups to discern potential sources of differences.

Misuse of the three options of encoding third person referents (NS, LS, OSP) was predicted in the performance of bilinguals based on accounts of interface conditions, crosslinguistic influence from Spanish and findings of previous related studies. HS and L2ers were predicted to overuse overt subjects, an expectation which was confirmed only regarding LS. All groups of bilinguals were expected to overuse OSP in the two contexts, but this prediction was disconfirmed. Misuse of NS in TS was also expected to some degree in the bilingual production based on relevant findings in previous studies, this expectation being attested by the findings.

The first observation with regard to the bilingual speakers' production was the lower mean length of narratives by the HS and L2ers compared to Greek monolinguals and immigrants (see §6.2.1). In other words, the groups composed of speakers whose language of narration was their weaker language produced shorter pieces of discourse compared to the groups of speakers who narrated in their stronger language, i.e. monolinguals and immigrants. Indeed, as Montrul (2008: 247) observes, HS and L2ers share important characteristics including 'reduced and variable exposure to the target language and degree of motivation needed to seek opportunities to use the language'. In the present study, both these groups of speakers were dominant bilinguals displaying greater ease in Spanish. The other type of bilinguals, i.e. the immigrants, were monolingually raised and schooled in Greek; thus, although using predominantly Spanish in daily life, for them Greek was more dominant than for HS and L2ers. The fact that the monolingually raised groups outperformed the HS and L2ers in narrative length was expected since the latter exhibited different degrees of command in Greek. Speaking in a weaker language entails being generally less fluent. Analogous findings have been reported in other studies. For instance, Andreou (2015: 254) explains similar differences found in narratives' length between monolingual and bilingual children as 'possible differences in lexical retrieval and possibly increased competition between languages'. Nonetheless, in other studies this difference was not attested, such as in Polinsky (2008: 151), in which Russian HS and monolinguals produced narratives which were relatively similar in length since 'they followed the same story plot'.

With regard to the type of clauses used by bilinguals, the three groups along with Greek monolinguals produced similar rates of matrix and embedded clauses (see §6.2.2). This contrasts with what has been reported in the literature with particular regard to HS, namely that they may tend to avoid embedded clauses (Polinsky 2008). The relative frequency of embedded clauses was associated with proficiency in both groups of HS and L2ers, with less proficient speakers producing fewer embedded clauses than more advanced speakers (§6.2.2).

The overall pattern was relatively similar in the bilingual groups in their Greek performance revealing sensitivity to the pragmatic conditions of subject distribution.

The L2 group produced more LS and fewer NS compared to the others. In all contexts, NS were used in more than half the time by the bilinguals and were the subject form favoured in embedded clauses, as in monolinguals. The relative frequency of LS was 32% in immigrants and 36% in HS, being close to Greek monolinguals (33%), but it was significantly higher in the L2 group, i.e. 41.5%. The OSP were scarcely produced by the three bilingual groups (0.3-1.2%), similarly to monolingual Greek. The most frequently used type of embedded clause was the adverbial, comparably to monolinguals' production. Subject distribution was examined more thoroughly within the particular contexts of use.

6.6.2.1 Topic continuity

TC contexts were expressed through the use of NS most of the time in the three bilingual groups, while LS were also produced in this context to a much lesser extent, similarly to monolinguals. Significant differences were found between monolingual and immigrant Greek on the one side and HS and L2ers on the other respecting the use of LS in TC, which was higher in the latter groups relative to the former. Likewise, similar differences were found on the use of NS, which was higher in Greek monolinguals/immigrants and lower in HS/L2ers. Production of OSP was minimal and in all cases conveyed focus. An exception was one low-proficiency L2 speaker who was found to overuse OSP in TC.

Lexical subjects

Statistical analyses confirmed that the L2 and HS groups used LS in TC significantly more often than Greek monolinguals and immigrants (§6.3.1.1). The qualitative analysis showed that most of the uses of LS in TC in the three groups were not redundant, while the frequences of redundant LS were not significantly different between monolingual and bilingual groups (§6.3.1.2). The potential reasons for production of LS in TC, which at first blush appears to be totally infelicitous, are the same was those stated in §6.6.1.1 regarding monolinguals. What needs to be explained is the quantitative difference between LS in TC in the groups of HS and L2ers vs Greek monolinguals and immigrants.

The relative frequency of LS in TC was 7.5% in HS and 12.3% in L2ers, i.e. significantly higher in the latter group. The basic difference between these groups compared to Greek monolinguals and immigrants is, as already mentioned, the fact that the latter speakers were monolingually raised and schooled in Greek. They were, thus, more dominant and more proficient in Greek than HS and L2ers.

Similar results were reported in Hendriks (2003), who explains the phenomenon of over-explicitation in reference maintenance in the context of L2 acquisition as reflecting the effort of speakers when constructing discourse in minimising redundancy and opacity at the same time. The felicitous pattern requires a balance between the desire of economy and the need for sufficient information (Levinson 2000). Hendriks (2003: 294) argues that 'the search for clarity' may seem pronounced in adult L2ers 'given that they are aware of their "short-comings" in the L2'; hence, over-explicitation 'has to do with a concern for communicative success' (ibid: 294). In the same vein, Bini (1993) maintains that learners are conscious of their limited command of L2 structures, thus they prefer using overt forms so that accuracy in verbal morphology is not required for reference disambiguation. Leclercq & Lenart (2013) report that even advanced L2 learners preferred to avoid risks by choosing non-ambiguous means, such as LS for reference maintenance, even at the expense of discursive cohesion. They argue that 'the learner is faced with a cognitive overload due to the incomplete mastery of the linguistic system of L2. This overload constrains discourse planning', leading learners to select more specific referential forms, 'which provide semantic content without relying on the interlocutor's retrieval of information from context' (ibid: 27). Such claims regarding L2ers can be easily extended to HS. Although the two types of bilinguals had different exposure to Greek input, both speak Greek as their weaker language.

Lozano (2018) similarly asserts that producing informationally richer phrases than pragmatically required is a characteristic of L2 learners. This may be interpreted as 'a consequence of enhanced perspective-taking abilities' (Sorace 2016: 6; see also Hendriks et al. 2014). Such activation of superfluous information may however result in processing cost (Gelormini-Lezama & Almor 2011). In any case, as Arnold et al. (2009: 146) hold, 'overspecified references may sound clumsy or pedantic, but

ultimately result in the listener successfully understanding the message'. Being uneconomical in subject reference often concerns bilinguals' overuse of OSP in TC in the literature (e.g. Bini 1993; Sorace et al. 2009; Sorace 2016; Lozano 2018). In the context of the present study, however, the phenomenon of subject redundancy in TC involving overt forms emerged only with LS and not with OSP.

Overt subject pronouns

Against the initial predictions, there was no overuse of OSP in TC contexts in any of the bilingual groups. This finding leads to the assertion that predictions stemming from the Interface Hypothesis (IH) were disproved. In fact, OSP were produced sparingly by bilinguals, similarly to Greek monolinguals. The few occurrences of OSP in TC involved focus, except for the L2 group where one low-proficiency speaker produced four OSP in TC, three of which were totally redundant. The fact that the speaker who produced infelicitous instances of OSP was a less proficient speaker indicates that such use was made to compensate for disfluency (e.g. hesitation, filling pauses) and poor command of morphology, along the lines of Bini (1993). As shown in several studies (e.g. Bini 1993; Pérez-Leroux & Glass 1997; Margaza & Bel 2006; Lozano 2006a, 2018), overuse of OSP in TC is diminished as higher levels of proficiency are achieved. Crucially, no misuse of OSP was found at any stage of potential L1 attrition or in near-native speakers of Greek, hence no signs of OSP vulnerability emerged as claimed by the IH. This is important bearing in mind that IH makes claims on interface vulnerability particularly with regard to final stages of L2 acquisition (near-natives) and initial stages of L1 attrition.

The present study does not corroborate previous findings regarding overextension of the scope of OSP in TC contexts in production of bilinguals of two NS languages. According to Sorace (2011, 2012), due to bilinguals' allocation of attentional resources, compensation of processing cost is accomplished by resorting to default strategies, with (over)use of OSP as the default being recurring even in pairs of two NS languages (see §3.4). The present data seem to support the claim that the default in TC production is more likely to be the LS, which was the overused form in the performance of HS and L2ers, and not the OSP. This, however, could be an artifact of

the narratives task that introduced more than two characters. It is worth reiterating that differences between the HS/L2ers and Greek monolinguals/immigrants in LS use were only quantitative, thus there was no actual pragmatic loss in subject use.

The question emerging is why the OSP was not overused in TC by bilinguals as in previous studies. The first observation is that in Greek (and Spanish) OSP are more often omitted than expressed. This may have to do with the nature of the third-person OSP in Greek, which is relatively rare in production by virtue of its strong deictic nature, hence infrequent in the input that HS and L2ers receive. Relative scarcity of OSP implies that positive evidence in the input is limited. Low use of OSP by bilinguals may thus reflect sensitivity to evidence in the input in terms of frequency effects. It also indicates that production of OSP in Greek is not vulnerable in bilingual situations in contact with another NS language, such as Spanish. The strong nature of the Greek OSP bearing the discourse-interpretable features of [+TS] and [+Focus], thereby appearing almost always in TS and focused contexts, possibly makes it resistant to variation in production of TC contexts. Since the OSP is more constrained, it may be simpler to detect its discourse behaviour in the input and at the same time its relatively categorical distribution makes it impermeable to contact effects when the other language is not substantially different. The prediction of the representational account (see §3.2) is born out by these findings. The two languages instantiate a complex pragmatic setting, thus no underspecification of the discourse-interpretable features would be expected, hence no optionality in the performance of bilinguals of this language pair (see Tsimpli et al. 2004; Sorace & Serratrice 2009).

HS and L2ers narrating in their weaker language, however, tended to be overspecific by favouring LS in more contexts than Greek monolinguals and immigrants. In the production of bilinguals, therefore, LS as marked anaphoric expressions could be seen as playing the role of the referential default used to mitigate processing overload. This may be due to the idiosyncrasy of the Greek OSP, whose use is more restrained and unaffected by language contact. Thus, bilinguals are led to access alternative default structures, with LS being a good candidate by virtue of its informativeness and immediate availability.

The age factor

An association between age and use of LS over NS in TC was found in the L2 group in which, similarly to Greek monolinguals, being older decreases the chance of using LS over NS. In other words, the younger the L2ers the more explicit were their referring expressions. This could be explained similarly to what was stated in §6.6.1.1 with regard to monolinguals. In addition, as L2ers have learned the language through instruction, they were familiar with following guidelines to complete language-related tasks. Adherence to the instructions of the task may have partly caused over-informativeness (use of LS in TC) in younger L2ers, while their older peers preferred the most economical option, i.e. use of NS.

6.6.2.2 Topic shift

The general distribution of subjects in TS contexts was relatively similar in the three bilingual groups and comparable to monolinguals. Namely, LS were predominantly used in TS with a relative frequency of 70% in immigrants, 63.5% in HS and 77% in L2ers, the latter group exhibiting the highest percentage of all. NS were the next subject form preferred in TS, with a range of occurrence of 29% in immigrants, 35% in HS (the highest rate) and 21% in L2ers. The instances of OSP were few, similarly to Greek monolinguals. The main statistically significant differences arising from the data is the high rate of LS in TS found in L2ers and the high rate of NS in TS detected in HS. The immigrants' performance was overall similar to Greek monolinguals and HS, lying in between these two groups with respect to subjects rates in TS.

Lexical subjects

LS were favoured in TS contexts involving reintroduction of a referent, in line with previous studies (e.g. Lozano 2009, 2018; Leclercq & Lenart 2013; Hendriks et al. 2014; Tsimpli et al. 2014). Lozano (2009) observes that in TS contexts the default option seems to be the full NP, which was found to encode TS more frequently than OSP in written corpora of native and non-native Spanish. The L2ers of the present study used a significantly higher rate of LS in TS relative to HS. Correspondingly, HS

used a significantly higher percentage of NS in TS relative to Greek monolinguals and L2ers (see §6.3.2.1).

The question is why L2ers were more informative in TS in producing more LS than HS, while HS were underexplicit in overusing NS compared to L2ers. Given that in TS contexts the felicitous option is the overt subject, the problem seems to lean towards HS' under-informative performance in TS. Overuse of NS in TS by HS, as well as the ambiguity potentially stemming thereof, are the focus of following sections. As regards the relatively high rate of LS in TS in the L2 group, these speakers may use more widely the LS as the default referential form in TS. More frequent use of LS in TS may additionally demonstrate the attempt of L2ers to be adequately informative without risks of miscommunication, similarly to TC contexts as explained previously. This also potentially reflects their familiarity with task completion following guidelines (i.e. clearly referring to the plot as instructed) since their learning of L2 Greek mainly (or at least initially) took place in an instructed environment.

Overt subject pronouns

OSP in TS were used by the three bilingual groups in rates which were similar to that of Greek monolinguals, i.e. in no more than 2% of the contexts. The qualitative analysis showed a slight qualitative difference between monolingual Greek vs bilingual Greek regarding OSP production in TS. Namely, in monolingual Greek, the few OSP used in TS were required in the context for subject/referent disambiguation, while in the immigrants and HS' performance some of the few OSP in TS could be omitted without causing referential ambiguity. In the L2 performance, the small number of OSP in TS was produced by speakers of basic proficiency, as in TC. Nonetheless, there were no pronounced differences between the Greek-speaking groups and the relative frequency of OSP was particularly low.

Crucially, since the crosslinguistic difference resided in the distribution of OSP in TS contexts, influence from Spanish was expected to some extent in the Greek performance of bilinguals due to language contact and Spanish being the dominant language. However, no crosslinguistic influence was found since the bilinguals did

not overuse OSP, suggesting that the Greek OSP is solid and invulnerable to influence from Spanish. This may be attributed to the nature of the Greek OSP and to the fact that, although statistically significant differences were obtained in the rates of OSP produced in TS, the two languages were not drastically different. Put differently, Greek and Spanish OSP distribution is not sufficiently different in order to trigger crosslinguistic influence. At least in production, Greek and Spanish present a rather low-level variation with respect to OSP, i.e. the differences are relatively small. This indicates variation at a microparametric level applying in the context of pronouns (Biberauer & Roberts 2012) and in language-specific micro-cues which specify the relevant contexts for small parts of the grammar in the acquisition process (Westergaard 2014). Thus, interface vulnerability was not shown to apply to Greek OSP in TS either. This finding also points to the relevance of the language combination for crosslinguistic influence.

Null subjects

NS were used in TS contexts by the three bilingual groups in generally comparable rates to those of monolinguals. Similar results were reported in Pinto (2014). However, HS were found to use significantly more NS in TS than Greek monolinguals and L2ers. The quantitative difference in the overuse of NS in TS implies overgeneralisation of the scope of the unmarked form by the HS. These speakers were thus found to be relatively overexplicit in TC but underexplicit in TS. HS were simultaneous bilinguals of two NS languages, with Greek as their weaker language naturalistically acquired through oral input. The reasons behind overuse of NS may relate to input properties, (over)use of NS as the referential default and language-external cognitive resources, as expounded below. In any case, this use seems to be driven by economy principles and it cannot be attributed to crosslinguistic influence.

Observably, NS are very frequent in NS languages, more so than LS or OSP, corresponding to topics, i.e. encoding continuity of old/given information (Dimitriadis 1996; Pérez-Leroux & Glass 1999; Georgopoulos 2017). Even in TS, NS are more common than OSP as revealed in the present findings (see also Shin & Cairns 2012; Pinto 2014). It follows that, in detecting probabilistic patterns, evidence in the input is

ambiguous since the rules seem to be flexible. Although HS overused NS in TS compared to Greek monolinguals and L2ers, their performance was not different from that of immigrants, who are presumably their main source of input (Montrul & Polinsky 2011; Kaltsa et al. 2015; Montrul 2016b). Crucially, older immigrants used NS in TS at very similar rates to HS (see §6.4.3). Thus, apart from NS being the most common subject form in the input, older immigrants, who are the closest we can get to the HS' parental generation, manifested additional frequency of NS in production. Overuse of NS in TS in HS may therefore relate to evidence in the input[16], since HS may receive qualitatively different input with occasional misuse of null subjects (i.e. emerging optionality) (Montrul 2016a; see also Sorace 2005). This could be a sign of intergenerational attrition (see Pascual y Cabo & Rothman 2012).

Overuse of NS in TS was found in HS and older immigrants but not in other older participants. Consequently, this phenomenon is primarily related to bilingualism in reduced input conditions, with disuse and the degree of language contact affecting the general cognitive ability in production. In such case, it is the LoR of older immigrants which matters rather than their age at testing, since being long-term immigrant implies higher language-contact intensity. Thus, HS being simultaneous bilinguals and older immigrants with long LoR, i.e. groups of speakers with longer experiences of bilingualism, manifest language-contact effects surfacing as overgeneralisation of NS in TS. Reducing use of structures is potentially due to constraints in the bilingual performance system or working memory effects, with subject structures tending to be dropped more than other elements because of their usually initial position (Rizzi 2000). The phenomenon may be thus a consequence of taxed processing resources manifested as a weakness at the level of language control (on-line processing) (see Sharwood Smith & Van Buren 1991; Flores 2008), especially when it leads to ambiguous constructions. It may also be an indication of loss of pragmatic content in cases of ambiguity, as will be discussed in §8.2.

---

[16] Although the immigrants acquired Greek in a monolingual context in Greece, the Greek input that they receive in Chile partly also comes from HS. This holds, of course, for speakers belonging to families with heritage-language-speaking offspring and/or close relatives. Such cases suggest that HS can be also viewed as input-givers.

Due to cognitive constraints caused by bilingualism combined with reduced exposure to Greek input and disuse, HS/older immigrants apparently tend to overextend the linguistic default of the language (NS) in more demanding contexts such as in subject/topic discontinuity (TS). The linguistic default (the weaker element) contrasts with the 'learner default' (Tsimpli 2011: 101), which is the unfocused OSP in TC. In the present study, HS/older immigrants manifested a preference for the linguistic default (NS) while L2ers used the LS (not the OSP) as a potential form of default in TS. Overuse of the linguistic default may have to do with general cognitive constraints due to bilingualism, which cause difficulty in mapping syntactic and pragmatic conditions in production. As Gelormini and Almor (2011) note, NS require less articulatory effort than overt forms, thereby involving a more economical option, hence a less demanding cognitive task. Moreover, NS are largely felicitous in TS provided that clarity of reference is not at stake. In fact, as Clements and Domínguez (2016: 24) argue, NS 'are good candidates for default forms since their morphosyntactic content can be recovered by other means such as verbal morphology'. The bilingualism effect is arguably increased in HS/older immigrants compared to L2ers due to length of language contact in reduced input conditions. Thus, the former speakers resort to the linguistic default due to taxed processing resources, while the latter favour more conservative strategies in using the marked referring expression (LS) and being more informative. The reason behind the fact that HS were relatively overexplicit in TC yet underexplicit in TS may lie in the fact that it is cognitively easier to repeat a recently mentioned subject in straightforward contexts (TC) than to produce a different subject form in more complex contexts (TS).

This difference between HS and L2ers may also concern their acquisition trajectory. HS are naturalistic language acquirers, who are normally better in oral modality and have also reduced contact with formal language registers. In contrast, L2ers as previous classroom learners are more 'artificial' speakers, who often perform better in written modality. In this regard, L2ers are regarded as 'hyperliterate' (Montrul 2014: 293). It is common knowledge that in written modality there is a tendency for enhanced explicitness compared to spoken language. In the latter case, speakers assume mutual knowledge more easily than in writing due to pragmatic principles

(e.g. principle of least effort, cooperative principle); hence, they can omit information more liberally. HS as oral acquirers thus tend to be less explicit because they are used to the oral modality involving colloquial registers, which are often underexplicit compared to written language. For L2ers, having learned the language in a more structured environment may be a contributing factor to the increased awareness of ambiguity, which leads to 'safer' pragmatic referential strategies (see also Montrul 2016b).

Overproduction of NS in environments where overt subjects would be discursively expected has been attested in previous research on bilingual speakers (e.g. Montrul 2004a; Montrul & Rodríguez Louro 2006; Rothman 2009a; Shin & Cairns 2012; Montrul & Sánchez-Walker 2015; Montrul 2016a). The crucial condition is that NS do not give rise to inaccurate or incomplete interpretations, i.e. to referential ambiguity, as this may lead to 'communicative breakdown' (Lozano 2018: 427).

Ambiguity

Givón (2016: 29) asserts that 'in spontaneous spoken language, the mental proposition often appears as an elliptic, truncated structure, with zeroed out arguments', since speakers rely on context and human inferential abilities to fill in the missing bits. Thus, utterances with missing overt subjects are not uncommon, but crucially the expected interpretation of reference should be consistent with the interlocutors' state of mutual knowledge (Blackwell 2003) complying with the degree of activation or salience of the referent (e.g. Ariel 1990; Almor 2000). NS in TS sometimes resulted in ambiguity in the production of bilingual speakers. According to Winter-Froemel and Zirker (2015: 283), 'ambiguity is frequently regarded as a "weak point" in semiotic systems, as it potentially leads to misunderstandings or, at least, increased processing efforts'. Since ambiguity may inhibit communication, it is deemed to be more problematic than redundancy in subject use.

The participants were invited to tell the stories to a hypothetical listener who could not see the pictures (see §5.2.4.1). The referents were thus supposed to be absent from the visual context for the listener and be present only in the shared prior

discourse context as established by the speakers. It was assumed that they should aim at avoiding under-explicitness by using unambiguous forms in order to lead to proper recoverability of subject referents, hence the correct meaning. This is generally expected to be enhanced when the discourse involves two or more characters (Arnold & Griffin 2007), as in the case of the stories used in this study. In view of absence of an explicit subject, if the set of possible referents is not reduced to the intended one via grammatical and/or contextual means, the listener may select an interpretation which is incongruent with the speakers' intended reference.

The three bilingual groups produced a number of ambiguity instances in contexts of TS involving NS, unveiling a problem of reference underspecification related to low degrees of sensitivity to antecedent accessibilty. Occurrence of ambiguity was found in the bilingual data while it was negligible in the monolingual data. Ambiguity was temporary or permanent (full), with the latter being arguably more problematic since it generally remains unresolved or definitely misguides listeners (see §6.3.2.3).

Permanent/full ambiguity was generated by two different causes. The first cause was the defective realisation of overt morphology and not the NS per se. The fact that such cases were also attested in the performance of advanced/near-native speakers indicates that it was not a mere developmental problem. Such occurrences point to the morphology-semantics/syntax as being 'the bottleneck' in bilingual production, which posed challenges in using functional morphemes and their features, along the lines of the Bottleneck Hypothesis (Slabakova 2008; see also Lardiere 2011). Similar claims were made by Zombolou (2011: 25) on Argentina-Greek: 'Greek is a language with very rich morphology, a fact that causes both the monolingual and the bilingual acquirers to confuse forms and mappings on grammatical structures'. Pelc (2001: 44) also observes with regard to Greek-English bilinguals (attriters) in the US context that Greek places 'a similarly heavy memory and processing load on their users due to the complexity of their noun paradigm and strict agreement requirements'. Thus, it is common that such grammar-internal interface gives rise to 'residual morphological variability' (Lardiere 2011: 51). In this respect, Rothman (2007b) notes that problems with grammatical morphology do not necessarily reflect competence since real-time spontaneous production may underdetermine the actual linguistic system and lead to

misinterpretations. This claim holds for both L1 and L2 performance, although it was raised regarding the L2 context, and points to judicious evaluation of speakers' production before resorting to claims about systematic differences.

The second cause of full ambiguity was missing overt subjects (genuine ambiguity). In environments lacking adequate cues for correct identification, an underspecified expression, i.e. a NS, was chosen when the intended referent was not a sufficiently salient entity to match the expression. Similar findings on ambiguity due to NS in TS have been reported in previous studies (Montrul 2004a; Montrul & Rodríguez Louro 2006; Miltsakaki 2007; Montrul & Sánchez-Walker 2015; Montrul 2016a). Since the participants did not rely on short-term memory to produce the narratives, as Hendriks et al. (2014) argue on similar work, imperfect story recall is eliminated as a confounding factor. Instead, such instances indicate an inefficient ability to integrate changing grammatical and pragmatic information by assessing and updating the discourse context in relation to the interlocutor's knowledge state. Some bilinguals thus apparently had more relaxed production abilities, which affected performance in more demanding contexts, i.e. TS, revealing 'egocentric' computations.

Age was significantly associated with ambiguity, i.e. the older the speakers the greater the chance of producing ambiguous NS, and this was particularly relevant in immigrants, but crucially it was not found in monolinguals. Hendriks et al. (2014) note that elderly adults differ from young adults in their productive referential skills due to limitations of memory span, which sharply increase from 75 years old onwards (see also Abrams & Farrell 2010; Shafto & Tyler 2014). In the present study, all cases of genuine ambiguity produced by immigrants involved speakers aged 76 or older. Ambiguous reference in older speakers, according to Hendriks et al. (2014), concerns the inefficient ability to take the perspective of the conversational partner. This is an extra step which requires additional cognitive resources and also depends on processing speed (see also Shin & Cairns 2012 on children's discourse).

The extended use of NS in TS in older bilinguals' production, sometimes resulting in ambiguity, seems to be also related to activation and inhibition mechanisms responsible for the control of two competing languages in the brain (Paradis 2004). It

is worth mentioning here the higher frequency of NS in TC found in the older Greek monolinguals and L2ers as compared to their younger counterparts, which is also an indication of preference for reduced forms by older speakers. According to Abrams and Farrell (2010: 50), the inhibitory processes which regulate the information that 'enters and leaves working memory' are weakened in older speakers in general. Bilingual memory of older speakers thus seems to be doubly affected by both the process of inhibition related to bilingualism and the age-related inhibition problems influencing language production. The reason why the age effect involving NS in TS and ambiguity was not found in monolinguals may be partly the fact that the Greek monolingual group did not include speakers of very advanced age. Miltsakaki (2007) also reports some cases of ambiguity in written production data of Greek monolinguals, although there is no mention of the age of the participants.

Emergence of ambiguity was unexpected in L2ers since these speakers tend to be rather overexplicit. A close look at the L2 participants who produced such ambiguities showed that half of them (N=3) were low-proficiency speakers. The other half (N=3) were advanced/near-native speakers who had lived in Greece for long periods of time, hence having been exposed to oral naturalistic input more than typical L2ers. Thus, in the L2 context one factor contributing to ambiguity in discourse is proficiency level, similarly to what was previously stated on command of morphology. The other factor is substantial exposure to oral colloquial language, which is less explicit than written language. As already mentioned, L2 learners/speakers are predominantly exposed to the written modality (at least more than HS), a fact which arguably leads them to be more informative when expressing reference. A number of L2ers had been exposed to naturalistic input significantly more than others. Some of these speakers thus manifested instances of underexplicitness causing ambiguity despite their advanced/near-native command of the language. This was potentially related to substantial exposure to oral language.

Noticeably, ambiguity involved TS contexts, while TC contexts were non-problematic as to reference. TS was found to be more complex due to its intrinsic unpredictability since, due to shift, the referent becomes less accessible (Givón 1983, 2016; Ariel 1990; Paredes Silva 1993) considering that shared knowledge of the storylines was not

assumed. This involves the fact that 'reintroducing a referent after a topic shift crucially requires speakers to take into account the listener's perspective' (Hendriks et al. 2014: 395). Older speakers resort to more economical options when failing to determine the discourse prominence of the referent because they encounter difficulties in 'keeping track of the structure of the discourse' (ibid: 404). In bilinguals, the processing cost leads to inconsistent efficiency in integrating syntactic and pragmatic information (Sorace 2011, 2016). It seems to be the case that a NS sometimes functions as the default even when it is not the optimal possibility, because it entails cognitive economy. In such cases, bilinguals do not optimise from the perspective of their listener since perspective taking is cognitively demanding.

However, indeterminacy of the optimal referential choice may occasionally lead bilinguals to simply omit an overt form in real-time production. This may wrongly appear as a deliberate use of NS, without reflecting a genuine grammatical choice (Clements & Domínguez 2016). Such a misuse of NS could generate more problems in communication than overusing an OSP. The fact that ambiguity was related to age but it was minimally found in monolinguals indicates that it is a language-contact outcome, not exclusively related to, but rather enhanced by bilingualism.

7. Study 2: Anaphora Resolution

The present chapter concerns the interpretation task on anaphora resolution (AR). The research questions and predictions are formulated in Section 7.1. Then, Section 7.2 offers an overview of the group results obtained from the experiment. Section 7.3 provides the regression analyses of the data. Section 7.4 presents a follow-up experiment on the resolution of the demonstrative pronoun in Spanish. Section 7.5 offers a summary and Section 7.6 discusses the findings.

7.1 Research questions and predictions

Drawing on previous research evidence and theoretical accounts, the research questions and predictions guiding Study 2 are formulated in the following sections.

7.1.1 First research question: Greek and Spanish monolinguals

Is Greek different from Spanish in the intepretation of anaphoric third-person null and overt subjects? If so, in which contexts and what causes divergence?

What is the effect of age at testing on the interpretation of ambiguous anaphoric subjects by monolinguals?

Predictions

In interpreting non-biased ambiguous forward AR, Greek and Spanish monolingual speakers are expected to perform similarly in resolving NS according to the predictions of PAH, thereby NS being assigned to matrix subjects. Considering, however, the findings in Mastropavlou et al. (2014), NS in Greek may reveal indeterminate antecedent preferences. In this case, the same would hold for Spanish. The Greek OSP is expected to consistently pick matrix objects, in line with PAH. In Spanish, the OSP is expected to be either ambiguous or also attached to objects but less strongly so than in Greek. In other words, OSP are expected to show weaker

resolution patterns in Spanish than in Greek and NS are expected to trigger relatively similar preferences in the two languages.

Since indefiniteness diminishes antecedent prominence, when an indefinite object is involved in the structure, the resolution of the NS might be more strongly directed to subject antecedents and the resolution of the OSP may be (more) biased towards object antecedents. Considering the findings in Mastropavlou et al. (2014), the matrix object definiteness was not expected to be significantly associated with antecedent preferences in Greek and the same would hold for Spanish.

The two languages were predicted to differ in the resolution of the OSP, with Greek being more rigid and Spanish more flexible. No categorical judgments were expected to be triggered in these contexts in any of the monolingual groups and conditions.

The age within adulthood has been found to play a role in resolving ambiguous pronominal anaphora (see Hendriks et al. 2014; Ghaleh 2015; Kaltsa et al. 2015). Thus, there may be age effects in the resolution of overt and null subjects, with the older participants exhibiting weaker preferences that their younger counterparts. In line with Hendriks et al. (2014), both older and younger speakers are equally able to identify the shift in contexts of TS; therefore, it is predicted that in such contexts, i.e. when an OSP is present, younger and older monolinguals should not differ. Differences related to age effects may emerge in the resolution of NS, given that NS are more ambiguous.

7.1.2 Second research question: Greek-Spanish bilinguals

Is Greek in contact with Spanish different from monolingual Greek in the interpretation of third-person anaphoric null and overt subjects? If so, in which contexts and what causes divergence?

What is the effect of age at testing on the interpretation of ambiguous anaphoric subjects by bilinguals?

Predictions

In interpretation of non-biased ambiguous forward AR contexts, bilingual Greek is predicted to be more inconsistent than monolingual Greek due to language contact and reduced input conditions. The third-person OSP is expected to be affected because of (a) crosslinguistic influence from dominant Spanish; and (b) interface conditions related to the distribution of subject pronouns, resulting in processing cost and interpretation of OSP as the default in pragmatically less appropriate contexts (i.e. in TC). Bilinguals are generally expected to comply with the PAH, but to manifest weakened PAH strategies. Namely, OSP are expected to allow for subject antecedents and NS to pick object antecedents more often than in monolingual preferences.

As regards definiteness, similarly to Greek monolinguals, the bilinguals' resolution preferences were not expected to be substantially influenced by the (in)definite matrix object, in line with Mastropavlou et al. (2014). NS could be more often assigned to subjects and OSP to objects in cases of indefinite objects when compared to definite object conditions.

Since the age of adult bilingual speakers may affect resolution of ambiguous anaphora (Ghaleh 2015; Kaltsa et al. 2015), aging effects are expected in the performance of bilinguals in the AR task, with older bilinguals exhibiting weak preferences compared to younger bilinguals and to monolinguals. The potential age effect entails some cognitive decline which affects working memory and attention resources (see Abrams & Farrell 2010). Differences related to age should be more pronounced in bilinguals than in monolinguals because bilingualism may intensify the tendency to use economical processing strategies, such as 'recency of mention' to interpret the NS. The effect of age in the performance of the participants is examined performing logistic regression analyses.

7.2 Overview of the results

This section offers an overview of the interpretation results based on the data elicited from the participants' responses to the AR task. As already explained in Chapter 5, four conditions were included in the oral questionnaire. The matrix subject was definite in all conditions. It was followed by a definite or an indefinite object in the same clause and an embedded clause with a null or overt pronoun in subject position as shown below:

(i)	Definite matrix object - Null embedded subject	(DDN)
(ii)	Indefinite matrix object - Null embedded subject	(DIN)
(iii)	Definite matrix object - Overt embedded subject	(DDO)
(iv)	Indefinite matrix object - Overt embedded subject	(DIO)

After listening to each sentence, the participants were asked who did the action of the embedded verb and they had to orally choose the antecedent by linking the embedded subject to the matrix subject or to the (in)definite matrix object. All groups performed the AR task only in Greek, except for the Spanish monolinguals, who performed the task in Spanish. The results are based on the responses of 115 participants (for details see Chapter 5). Although the AR contexts were ambiguous, in line with the literature it is presumed that NS are more often associated with contexts of TC and OSP with contexts of TS. The sentences (89), (90), (91) and (92) in the following sections are examples of the experiment items in the two languages (see Appendix D).

7.2.1 Condition DDN

DDN: Definite Subject – Definite Object – Null Subject

(89) *I jaja filuse ti nosokoma otan iði evaze to palto tis.*

   La abuela besaba a la enfermera cuando ya se ponía el abrigo.

   'The old lady was kissing the nurse when [adverb] was putting on her coat.'

Figure 7.1 and Table 7.1 present the results of the DDN condition. Greek and Spanish monolinguals performed similarly in identifying the embedded pronoun in this condition [Pearson χ^2 (1, N=160)=0.025, p=0.874]. Specifically, their matching decisions for the embedded NS did not show a preference towards either the subject or the object of the matrix clause (Greek: 46.25% vs 53.75%; Spanish 47.5% vs 52.5% for subject and object preferences respectively).

A similar behaviour was manifested in the case of immigrants (47.79% vs 52.21). This means that in monolingual Spanish and Greek as well as in immigrant Greek resolving NS reference in this condition was at chance level with no significant difference between the groups. The L2 group showed a preference towards the matrix object (58.33%) but without reaching statistical significance, indicating again a chance level performance (z=-1.394, p=0.163). In HS, the matrix object was clearly preferred (63.04%) over the matrix subject (36.96%) and the difference between subject and object antecedent preferences was statistically significant (z=-2.420, p=0.016).

No further significant differences were found between the groups in DDN. In particular, Pearson Chi-square tests showed no significant difference between the Greek monolinguals and the HS' resolution pattern in this condition [Pearson χ^2 (1, N=172)=1.524, p=0.217].
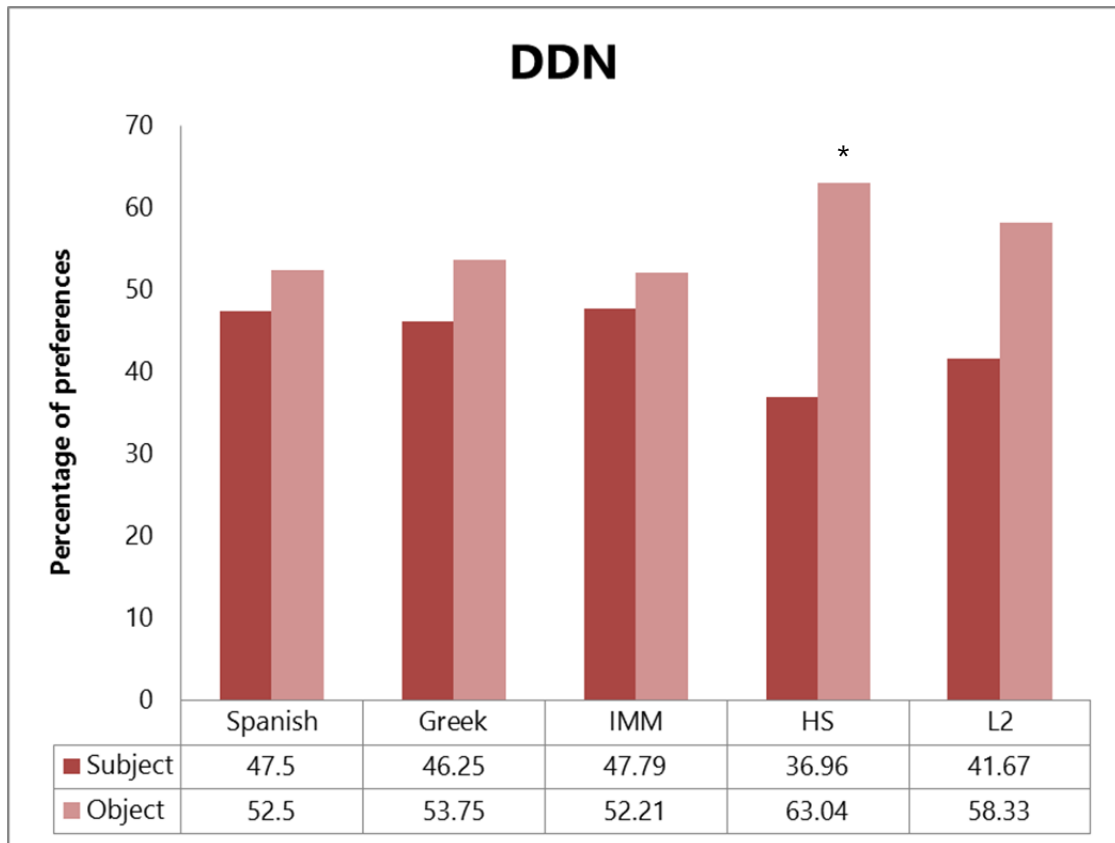
Figure 7.1. Group results in condition DDN (AR)

Table 7.1. Group results in condition DDN (AR)

	Spanish	Greek	IMM	HS	L2	Total
Subject	38	37	65	34	30	204
	47.5	*46.25*	*47.79*	*36.96*	*41.67*	*44.35*
Object	42	43	71	58	42	256
	52.5	*53.75*	*52.21*	*63.04*	*58.33*	*55.65*
Total	80	80	136	92	72	460
	100	*100*	*100*	*100*	*100*	*100*

7.2.2 Condition DIN

DIN: Definite Subject – Indefinite Object – Null Subject

(90) *I jaja filuse mia nosokoma otan iði evaze to palto tis.*
      La abuela besaba a una enfermera cuando ya se ponía el abrigo.
      'The old lady was kissing a nurse when [adverb] was putting on her coat.'

Figure 7.2 and Table 7.2 show the results of the DIN condition. The resolution routines in this condition were generally similar to those attested in the DDN condition. Greek and Spanish monolinguals exhibited a similar pattern with no bias towards either antecedent (Greek: 50% vs 50%; Spanish 46.25% vs 53.75% for subject and object preferences respectively). There was no statistical difference between the groups [Pearson χ^2 (1, N=160)=0.225, p=0.635].

The immigrants' resolution preferences were similar to their DDN responses (47.79% vs 52.21%) and similar to monolinguals. The L2 group showed again a preference towards the matrix object (54.17%), slightly lower than in DDN, but with no statistical difference between subject and object preferences (z=-0.705, p=0.480). In HS, the matrix object was preferred (58.7%) over the matrix subject (41.3%), but in this case there was no statistical difference between these preferences (z=-1.644, p=0.100), contrary to the DDN condition. No statistical difference was found in the responses of DDN and DIN in the HS performance [Pearson χ^2 (1, N=184)=0.365, p=0.546].

No clear preferences were therefore manifested in the DIN condition in any group performance regarding NS resolution. Thus, the AR patterns in DIN were at chance level with no significant difference between the groups.

Pearson Chi-square tests showed no significant differences between the groups of speakers in DIN or between the participants' DIN and DDN preferences. Indefiniteness of the matrix object did not appear to influence group preferences. However, definiteness seems to play a limited role in the preferences of HS, since in DIN they select the subject antecedent slightly more than in DDN.
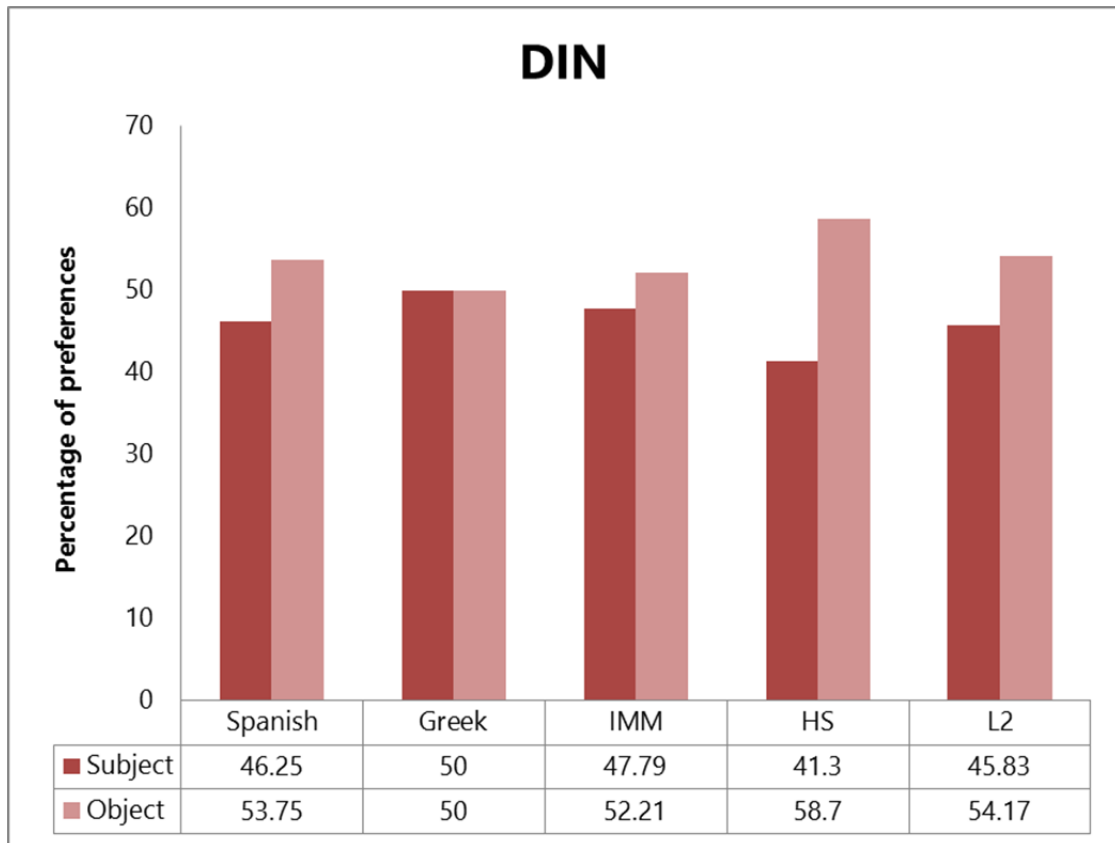
Figure 7.2. Group results in condition DIN (AR)

Table 7.2. Group results in condition DIN (AR)

	Spanish	Greek	IMM	HS	L2	Total
Subject	37	40	65	38	33	213
	46.25	*50*	*47.79*	*41.3*	*45.83*	*46.3*
Object	43	40	71	54	39	247
	53.75	*50*	*52.21*	*58.7*	*54.17*	*53.7*
Total	80	80	136	92	72	460
	100	100	100	100	100	100

7.2.3 Condition DDO

DDO: Definite Subject – Definite Object – Overt Subject

(91)   *I jaja filuse ti nosokoma otan afti evaze to palto tis.*

   La abuela besaba a la enfermera cuando ella se ponía el abrigo.

   'The old lady was kissing the nurse when she was putting on her coat.'

Figure 7.3 and Table 7.3 show the results of the DDO condition. A statistically significant difference was observed between Greek and Spanish monolinguals in the resolution of the OSP [Pearson χ^2 (1, N=160)=20.19, p<0.001]. While in Spanish there was no strong bias towards an antecedent (51.15% vs 48.75% for subject and object preferences respectively), Greek monolinguals consistently linked the OSP to the object (82.5%) revealing a significant difference in their preferences between subject and object antecedent (z=-4.874, p<0.001). In the Spanish group, preferences in DDO were similar to those in DDN [Pearson χ^2 (1, N=160)=0.225, p=0.635] and in DIN [Pearson χ^2 (1, N=160)=0.400, p=0.527]. By contrast, monolingual Greek preferences in DDO were significantly different from those attested in DDN [Pearson χ^2 (1, N=160)=15.22, p<0.001] and in DIN [Pearson χ^2 (1, N=160)=18.89, p<0.001].

Similarly to Greek monolinguals, the matrix object antecedent was significantly preferred over the matrix subject by the immigrants (84.56%), the HS (85.87%) and the L2 group (77.78%). In all the bilingual groups the preference for the object over the subject antecedent was found to be statistically significant (IMM: z=-6.631, p<0.001; HS: z=-5.591, p<0.001; L2: z=-4.121, p<0.001). The preferences of the immigrants were significantly different from their matching decisions in DDN and DIN [Pearson χ^2 (1, N=272)=32.92, p<0.001]. HS also significantly differed from their preferences in DDN [Pearson χ^2 (1, N=184)=12.60, p<0.001] and DIN [Pearson χ^2 (1, N=184)=16.95, p<0.001]. Finally, the L2 group also revealed a significantly different performance in DDO as compared to DDN [Pearson χ^2 (1, N=144)=6.26, p=0.012] and DIN [Pearson χ^2 (1, N=144)=32.92, p=0.003]. Pearson Chi-square tests showed no statistical differences between the Greek-speaking groups.
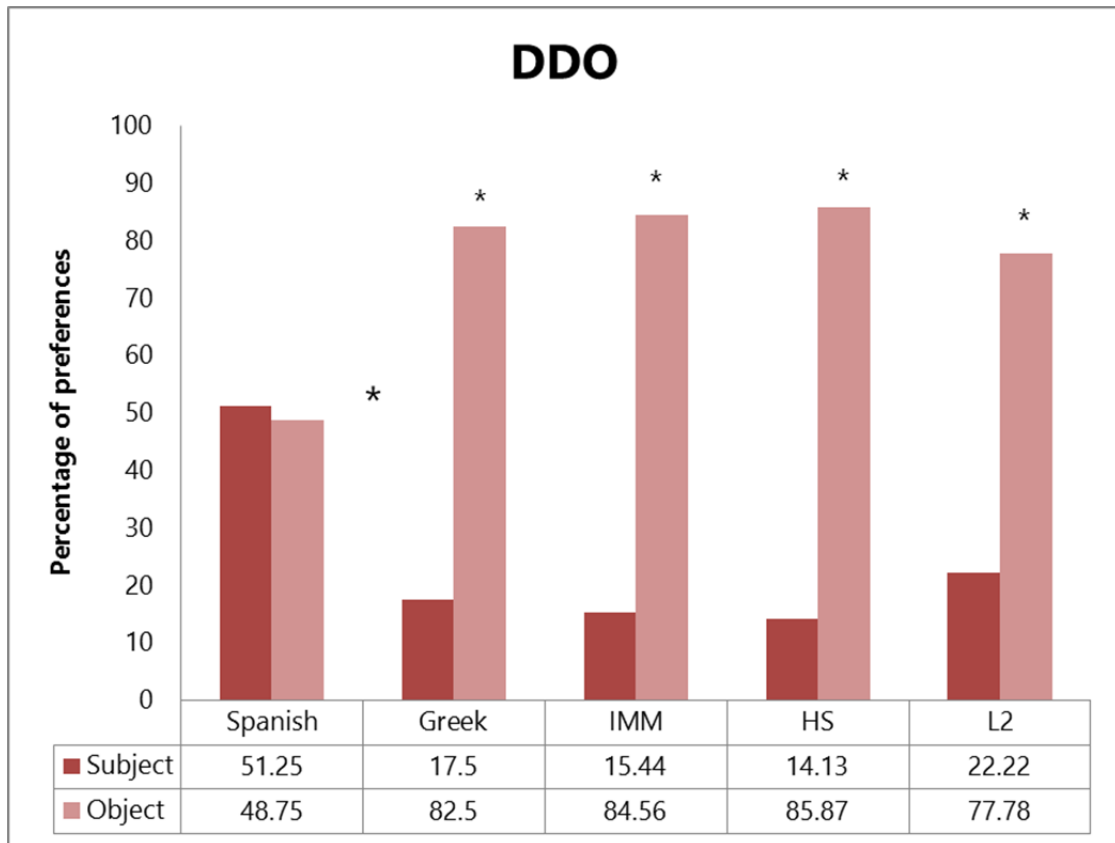
Figure 7.3. Group results in condition DDO (AR)

Table 7.3. Group results in condition DDO (AR)

	Spanish	Greek	IMM	HS	L2	Total
Subject	41	14	21	13	16	105
	51.25	*17.5*	*15.44*	*14.13*	*22.22*	*22.83*
Object	39	66	115	79	56	355
	48.75	*82.5*	*84.56*	*85.87*	*77.78*	*77.17*
Total	80	80	136	92	72	460
	100	*100*	*100*	*100*	*100*	*100*

7.2.4 Condition DIO

DIO: Definite Subject – Indefinite Object – Overt Subject

(92)   *I jaja filuse mia nosokoma otan afti evaze to palto tis.*

   La abuela besaba a una enfermera cuando ella se ponía el abrigo.

   'The old lady was kissing a nurse when she was putting on her coat.'

Figure 7.4 and Table 7.4 present the results of the DIO condition. In Spanish, there was a preference towards the matrix object, but the difference between subject and object preferences (41.25% vs 58.75% respectively) was not statistically significant (z=-1.541, p=0.123). A Pearson Chi-square test indicated no statistical difference between the DDO and DIO resolution patterns in Spanish [Pearson χ^2 (1, N=160)=1.609, p=0.205]. The Greek group significantly favoured the object (77.5%) (z=-4.310, p<0.001), although slightly less often than in DDO (82.5%), but with no statistical difference between DDO and DIO conditions [Pearson χ^2 (1, N=160)=0.625, p=0.429]. A statistical difference was found between Greek and Spanish monolinguals in resolving OSP in DIO [Pearson χ^2 (1, N=160)=6.47, p=0.011] with Greek monolinguals consistently picking the object antecedent and the Spanish performing at chance.

A bias towards the object antecedent was also found in the bilingual groups, with the immigrants selecting it 86.03% of the time and the HS and L2ers 75% of the time. As in DDO, the bilingual groups significantly favoured the object over the subject antecedent (IMM: z=-6.817, p<0.001; HS: z=-4.289, p<0.001; L2: z=-3.794, p<0.001). There were no statistical differences in preferences between DIO and DDO in the bilingual groups. Significant differences were detected between immigrants and HS [Pearson χ^2 (1, N=228)=4.44, p=0.035], as well as between immigrants and L2ers [Pearson χ^2 (1, N=208)=3.91, p=0.048], with the immigrants selecting the object more often than the latter groups. This may be due to age effects, since the group of immigrants included participants of older age. Age effects are explored in detail shortly. Indefiniteness of the matrix object did not appear to influence the participants' responses compared to the DDO condition.
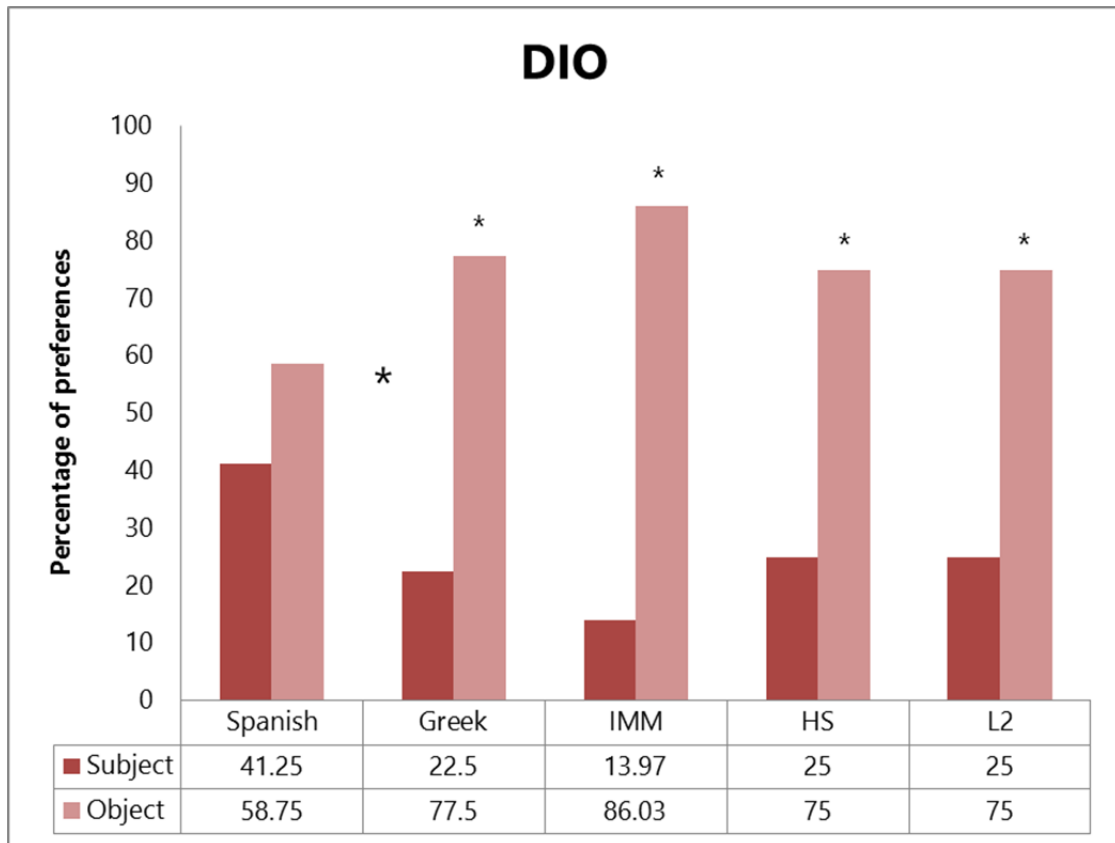
Figure 7.4. Group results in condition DIO (AR)

Table 7.4. Group results in condition DIO (AR)

	Spanish	Greek	IMM	HS	L2	Total
Subject	33	18	19	23	18	111
	41.25	*22.5*	*13.97*	*25*	*25*	*24.13*
Object	47	62	117	69	54	349
	58.75	*77.5*	*86.03*	*75*	*75*	*75.87*
Total	80	80	136	92	72	460
	100	*100*	*100*	*100*	*100*	*100*

7.3 Binomial logistic regressions

This section presents the results obtained from binomial logistic regressions conducted in the AR database. The analysis included the following three independent variables in their association with Antecedent Preference (AP) (subject or object):

(a) Group of speakers (Spanish monolinguals, Greek monolinguals, IMM, HS, L2)
(b) Age at testing (Age)
(c) Proficiency (only for HS and L2)

For purposes of consistency with the previous chapter, two sections are presented: TC and TS. There were thus two contexts with two conditions in each context:

A. TC with definite (DDN) and indefinite (DIN) object antecedent
B. TS with definite (DDO) and indefinite (DIO) object antecedent

The AP (subject, object) was used as the dependent variable. Sixteen statistical models were developed, with four models for each condition. The analyses examined TC (DDN, DIN) and TS (DDO, DIO) separately, corresponding to the models as follows:

(e) First model: analysis of Group of speakers and Age in association with AP in the monolingual groups (Greek, Spanish).

(f) Second model: analysis of Group of speakers and Age in association with AP in the Greek-speaking groups (Greek monolinguals, IMM, HS, L2) with Greek monolinguals as the baseline group.

(g) Third model: analysis of Group of speakers and Age in association with AP in the bilingual groups (IMM, HS, L2) with immigrants as the baseline group.

(h) Fourth model: analysis of Group of speakers, Age and Proficiency in association with AP in the groups of HS and L2ers.

Odds Ratio (OR), Standard Error of the OR (SE), z-value of the model, p-value of the model and 95% Confidence Interval (CI) of the OR are shown in the regression tables.

7.3.1 Topic continuity

7.3.1.1 Condition DDN

First model: Spanish and Greek monolingual groups

A binomial logistic regression was performed in order to analyse the variables Group of speakers and Age in the groups of Greek and Spanish monolinguals in association with AP with Spanish as the baseline group (N of observations: 160). The results showed a pseudo R^2=0.054 and a significant model (p=0.003). Age revealed a significant (p=0.001) and negative association with the subject AP (OR=0.971), indicating that being older by one year increases by 2.9% the chance of selecting the object antecedent in DDN. A follow-up binomial logistic regression showed that this association was significant only in Greek monolinguals (p=0.002, OR=0.958), not in the Spanish group (see §7.3.4). There was no significant association between monolingual groups and their AP (Table 7.5).

Table 7.5. Binomial logistic regression: AP in monolinguals in DDN

AP	OR	SE	z	p	OR CI	
					LB	UB
Age	0.971	0.009	-3.33	0.001	0.954	0.988
Group of speakers						
Greek	0.958	0.315	-0.13	n.s.	0.503	1.827

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Second model: Greek-speaking groups

A binomial logistic regression was conducted in order to analyse the variables Group of speakers and Age in the Greek-speaking groups (Greek monolinguals, IMM, HS, L2) in association with AP, with Greek monolinguals as the baseline (N of observations: 380). The results showed a pseudo R^2=0.062 and a significant model (p<0.001). Age showed a significant (p<0.001) and negative association with the subject AP (OR=0.969): being older by one year increases by 3.2% the chance of selecting the object in DDN. Further regressions indicated that the association between Age and AP was significant only in Greek monolinguals (p=0.002,

OR=0.958), as shown before, and the immigrants (p<0.001, OR=0.968) (see §7.3.4). There was no significant association between Greek-speaking groups and their AP (Table 7.6).

Table 7.6. Binomial logistic regression: AP in Greek speakers in DDN

AP	OR	SE	z	p	OR CI LB	OR CI UB
Age	0.969	0.006	-5.19	<0.001	0.958	0.981
Group of speakers						
IMM	1.269	0.378	0.80	n.s.	0.708	2.274
HS	0.828	0.268	-0.58	n.s.	0.439	1.564
L2	0.671	0.229	-1.17	n.s.	0.344	1.312

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Third model: Bilingual groups

A binomial logistic regression was conducted in order to analyse the variables Group of speakers and Age in the three bilingual groups (IMM, HS, L2) in association with AP, with immigrants as the baseline group (N of observations: 300). The results showed a pseudo R^2=0.052 and a significant model (p<0.001). Age showed a significant (p<0.001) and negative association with the subject AP (OR=0.972), indicating that being older by one year increases by 2.9% the chance of selecting the object antecedent in the DDN condition. A further binomial logistic regression showed that this association was significant only in the group of immigrants (p<0.001, OR=0.968), not in the HS and the L2 group (see §7.3.4). Additionally, no significant association was found between AP and Group of speakers (Table 7.7).

Table 7.7. Binomial logistic regression: AP in bilinguals in DDN

AP	OR	SE	z	p	OR CI LB	OR CI UB
Age	0.972	0.007	-4.19	<0.001	0.960	0.985
Group of speakers						
HS	0.650	0.185	-1.51	n.s.	0.372	1.136
L2	0.549	0.172	-1.90	n.s.	0.296	1.017

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Fourth model: HS and L2ers

A binomial logistic regression was conducted in order to analyse the variables Group of speakers, Age and Proficiency in HS and L2ers in association with AP, with HS as the baseline group (N of observations: 164). The results showed a pseudo $R^2=0.02$ and a non-significant model. Age, Proficiency and Group of speakers (i.e. being HS or L2er) were not significantly associated with AP (Table 7.8). Age, however, showed a marginally non-significant (p=0.053) negative (OR=0.980) association with AP, suggesting that the older the speakers the more likely may be to choose the object over the subject antecedent.

Table 7.8. Binomial logistic regression: AP in HS and L2ers in DDN

AP	OR	SE	z	p	OR CI LB	UB
Age	0.980	0.011	-1.93	0.053	0.958	1.000
Proficiency	0.962	0.165	-0.23	n.s.	0.687	1.345
Group of speakers						
L2	0.927	0.330	0.21	n.s.	0.460	1.865

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

A one-way ANOVA showed that there was a significant association between Age and AP in DDN when the two groups were analysed as one group [F (1,162)=0.701, p=0.043]. This indicates that considering HS and L2ers together, being older increases the chance of selecting the object antecedent in this condition (see §7.3.4).

7.3.1.2 Condition DIN

First model: Spanish and Greek monolingual groups

A binomial logistic regression was performed in order to analyse the variables Group of speakers and Age in Greek and Spanish monolinguals in association with AP (N of observations: 160) with Spanish as the baseline group. The results showed a pseudo $R^2=0.003$ and a non-significant model. Age and Group of speakers did not show a significant association with AP in monolinguals in DIN (Table 7.9). A further regression examining Age with each group separately revealed a significant (p=0.042)

and negative association (OR=0.975) in Greek monolinguals between Age and AP. As in DDN, the association indicated that being older increases the chance of selecting the object in the DIN condition (see §7.3.4).

Table 7.9. Binomial logistic regression: AP in monolinguals in DIN

AP	OR	SE	z	p	OR CI	
					LB	UB
Age	0.995	0.008	-0.64	n.s.	0.979	1.011
Group of speakers						
Greek	1.165	0.369	0.48	n.s.	0.626	2.169

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Second model: Greek-speaking groups

A binomial logistic regression was conducted in order to analyse the variables Group of speakers and Age in the Greek-speaking groups (Greek monolinguals, IMM, HS, L2) in association with AP, with Greek monolinguals as the baseline (N of observations: 380). The results showed a pseudo R^2=0.046 and a significant model (p<0.001). Age showed a significant (p<0.001) and negative association with the subject AP (OR=0.973), indicating that being older by one year increases by 2.8% the chance of selecting the object antecedent in the DIN condition. A further binomial logistic regression showed that the association between Age and AP was relevant only in immigrants (p=0.001, OR=0.971) and Greek monolinguals (p=0.042, OR=0.975) as seen in the first model for DIN (see also §7.3.4). There was no significant association between AP and Groups of speakers (Table 7.10).

Table 7.10. Binomial logistic regression: AP in Greek speakers in DIN

AP	OR	SE	z	p	OR CI	
					LB	UB
Age	0.973	0.006	-4.63	<0.001	0.962	0.984
Group of speakers						
IMM	1.059	0.311	0.19	n.s.	0.596	1.882
HS	0.832	0.265	-0.58	n.s.	0.446	1.552
L2	0.703	0.237	-1.050	n.s.	0.363	1.360

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Third model: Bilingual groups

A binomial logistic regression was conducted in order to analyse the variables Group of speakers and Age in the bilingual groups (IMM, HS, L2) in association with AP, with immigrants being the baseline group (N of observations: 300). The results showed a pseudo $R^2=0.050$ and a significant model ($p<0.001$). Age showed a significant ($p<0.001$) and negative association with the preference for the subject (OR=0.973), indicating that being older by one year increases by 2.8% the chance of selecting the object antecedent in the DIN condition. This association was relevant only for the group of immigrants (p=0.001, OR=0.971), as in the second model for DIN. There was no significant association between AP and Group of speakers (Table 7.11).

Table 7.11. Binomial logistic regression: AP in bilinguals in DIN

AP	OR	SE	z	p	OR CI LB	UB
Age	0.973	0.006	-4.170	<0.001	0.960	0.985
Group of speakers						
HS	1.189	0.564	0.370	n.s.	0.470	3.013
L2	0.685	0.214	-1.210	n.s.	0.371	1.265

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Fourth model: HS and L2ers

A binomial logistic regression was conducted in order to analyse the variables Group of speakers, Age and Proficiency in the groups of HS and L2ers in association with AP, HS being the reference group (N of observations: 164). The results showed a pseudo $R^2=0.032$ and a non-significant model. The variables Proficiency and Group of speakers (i.e. being HS or L2) were not significantly associated with AP in DIN. However, Age showed a significant (p=0.021) and negative association with the subject AP (OR=0.974), indicating that being older by one year increases the chance of selecting the object antecedent in DIN by 2.7% (Table 7.12). Age was analysed in this condition for each of the two groups separately and the statistical models appeared to be non-significant.

Table 7.12. Binomial logistic regression: AP in HS and L2ers in DIN

| AP | OR | SE | z | p | OR CI | |
					LB	UB
Age	0.974	0.010	-2.30	0.021	0.954	0.996
Proficiency	1.146	0.195	0.80	n.s.	0.820	1.600
Group of speakers						
L2	0.857	0.304	-0.44	n.s.	0.426	1.718

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

A one-way ANOVA showed that there was significant association between Age and AP in the DIN condition when both groups were included in the analysis [$F_{(1,162)}=6.09$, p=0.015]. This indicates that considering HS and L2ers together, being older increases the chance of selecting the object antecedent.

7.3.2 Topic Shift

7.3.2.1 Condition DDO

First model: Spanish and Greek monolingual groups

A binomial logistic regression was performed in order to analyse the variables Group of speakers and Age in Greek and Spanish monolinguals in association with AP (N of observations: 160) with Spanish as the baseline group. The results showed a pseudo R^2=0.112 and a significant model (p<0.001). Group of speakers showed a significant (p<0.001) and negative association with the selection of the subject antecedent (OR=0.198): being Greek monolingual increases by 5 times the chance of selecting the object in DDO. Age did not show a significant association with AP in monolinguals (Table 7.13).

Table 7.13. Binomial logistic regression: AP in monolinguals in DDO

| AP | OR | SE | z | p | OR CI | |
					LB	UB
Age	0.986	0.009	-1.48	n.s.	0.968	1.005
Group of speakers						
Greek	0.198	0.074	-4.34	<0.001	0.095	0.411

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Second model: Greek-speaking groups

A binomial logistic regression was conducted in order to analyse the variables Group of speakers and Age in the Greek-speaking groups (Greek monolinguals, IMM, HS, L2) with Greek monolinguals as the baseline group, in association with AP (N of observations: 380). The results showed a pseudo R^2=0.010 and a non-significant model, indicating no significant associations between the variables (Table 7.14).

Table 7.14. Binomial logistic regression: AP in Greek speakers in DDO

| AP | OR | SE | z | p | OR CI | |
					LB	UB
Age	0.991	0.007	-1.15	n.s.	0.977	1.006
Group of speakers						
IMM	0.903	0.344	-0.27	n.s.	0.428	1.904
HS	0.824	0.349	-0.46	n.s.	0.359	1.890
L2	1.279	0.526	0.60	n.s.	0.571	2.865

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Third model: Bilingual groups

A binomial logistic regression was conducted in order to analyse the variables Group of speakers and Age in the bilingual groups (IMM, HS, L2) in association with AP, with immigrants as the baseline group (N of observations: 300). The results showed a pseudo R^2=0.008 and a non-significant model. This indicates no significant associations between the aforementioned variables with AP in the DDO condition (Table 7.15).

Table 7.15. Binomial logistic regression: AP in bilinguals in DDO

| AP | OR | SE | z | p | OR CI | |
					LB	UB
Age	0.997	0.008	-0.31	n.s.	0.981	1.014
Group of speakers						
HS	0.903	0.345	0.27	n.s.	0.427	1.910
L2	1.517	0.580	1.09	n.s.	0.715	3.213

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Fourth model: HS and L2ers

A binomial logistic regression was conducted in order to analyse the variables Group of speakers, Age and Proficiency in the groups of HS and L2ers in association with AP (N of observations: 164). The HS was the baseline group. The results showed a pseudo $R^2=0.017$ and a non-significant model. The variables Age, Proficiency and Group of speakers (i.e. being HS or L2er) were not significantly associated with AP in the DDO condition (Table 7.16).

Table 7.16. Binomial logistic regression: AP in HS and L2ers in DDO

AP	OR	SE	z	p	OR CI	
					LB	UB
Age	0.996	0.014	-0.33	n.s.	0.969	1.022
Proficiency	0.833	0.180	-0.84	n.s.	0.544	1.273
Group of speakers						
L2	1.656	0.754	1.11	n.s.	0.678	4.043

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

7.3.2.2 Condition DIO

First model: Spanish and Greek monolingual groups

A binomial logistic regression was performed in order to analyse the variables Group of speakers and Age in Greek and Spanish monolinguals in association with AP (N of observations: 160) with Spanish as the baseline. The results showed a pseudo $R^2=0.076$ and a significant model (p<0.001). Age revealed a significant (p=0.005) and negative association with the selection of subject antecedent (OR=0.973): being older by one year increases by 2.8% the chance of selecting the object antecedent in DIO. A follow-up binomial logistic regression showed that the association was significant for Greek monolinguals (p=0.006, OR=0.955), not for Spanish monolinguals (see §7.3.4).

The variable Group of speakers also showed a significant (p=0.011) and negative association with the selection of subject antecedent (OR=0.396), indicating that being Greek monolingual increases by 2.5 times the chance of selecting the object antecedent in DIO (Table 7.17).

Table 7.17. Binomial logistic regression: AP in monolinguals in DIO

AP	OR	SE	z	p	OR CI	
					LB	UB
Age	0.973	0.010	-2.84	0.005	0.954	0.991
Group of speakers						
Greek	0.396	0.143	-2.56	0.011	0.195	0.806

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Second model: Greek-speaking groups

A binomial logistic regression was conducted in order to analyse the variables Group of speakers and Age only in the Greek-speaking groups (Greek monolinguals, IMM, HS, L2) in association with AP, with Greek monolinguals as the baseline group (N of observations: 380). The results showed a pseudo R^2=0.016 and a non-significant model, indicating no significant associations between Age and Group of speakers with AP (Table 7.18). Further logistic regressions examining Age separately for each group showed a significant association between Age and AP in the Greek monolinguals (as in the previous model) and immigrants. The association was negative in Greek monolinguals (p=0.006, OR=0.955) but positive in immigrants (p=0.006, OR=1.033). This suggests that older immigrants tended to select the subject, whereas older Greek monolinguals tended to select the object antecedent (see §7.3.4). In both these groups, however, the object antecedent was strongly favoured.

Table 7.18. Binomial logistic regression: AP in Greek speakers in DIO

AP	OR	SE	z	p	OR CI	
					LB	UB
Age	0.997	0.007	-0.38	n.s	0.984	1.011
Group of speakers						
IMM	0.568	0.208	-1.54	n.s.	0.277	1.165
HS	1.169	0.425	0.43	n.s.	0.574	2.383
L2	1.129	0.434	0.32	n.s.	0.531	2.398

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Third model: Bilingual groups

A binomial logistic regression was conducted in order to analyse the variables Group of speakers and Age in the three bilingual groups (IMM, HS, L2) in association with AP, with immigrants as the baseline (N of observations: 300). The results showed a pseudo R^2=0.024 and a non-significant model. Although the model was not shown to be significant, significant associations emerged between Group of speakers and AP, indicating a positive association with the selection of the subject antecedent in HS (OR=2.065) and L2ers (OR=2.360). Accordingly, being HS or L2er increases the chance of selecting the subject antecedent in DIO by 2 times and 2.3 times respectively (Table 7.19). This is also supported by the results of Pearson Chi-square tests shown in §7.2.4, where significant differences were observed in the immigrants preferring the object more often than HS and L2ers. However, the level of association in the Pearson Chi-square tests was very weak (Cramer's V=-0.13 in both cases), which may explain why the present model was not shown to be significant.

Although Age did not show a significant association, further logistic regressions examining Age separately for each group showed a significant positive association between Age and AP in the immigrants (p=0.006, OR=1.033) (see §7.3.4).

Table 7.19. Binomial logistic regression: AP in bilinguals in DIO

| AP | OR | SE | z | p | OR CI | |
					LB	UB
Age	1.010	0.008	1.27	n.s.	0.994	1.026
Group of speakers						
HS	2.065	0.716	2.09	0.036	1.047	4.072
L2	2.360	0.916	2.22	0.027	1.104	5.049

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Fourth model: HS and L2ers

A binomial logistic regression was conducted in order to analyse the variables Group of speakers, Age and Proficiency in the groups of HS and L2ers in association with AP, with HS as the baseline (N of observations: 164). The results showed a pseudo

R^2=0.019 and a non-significant model. Age, Proficiency and Group of speakers were not associated with AP in the DIO condition in the HS and L2 groups (Table 7.20).

Table 7.20. Binomial logistic regression: AP in HS and L2ers in DIO

| AP | OR | SE | z | p | OR CI | |
					LB	UB
Age	0.983	0.012	-1.33	n.s.	0.960	1.018
Proficiency	0.761	0.149	-1.39	n.s.	0.518	1.117
Group of speakers						
L2	0.806	0.329	-0.53	n.s.	0.362	1.800

(AP=Antecedent Preferences, OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

A summary of the hitherto findings obtained from the statistical analyses is offered in the next section. Following that, Age is further explored in §7.3.4.

7.3.3 Interim findings

7.3.3.1 Resolution of null subjects

In resolving NS in DDN, all groups performed at chance level, except the HS who showed a clearer preference for the object antecedent. Logistic regressions did not detect an association between HS and AP, suggesting that HS' preference for the object was relatively weak. There were no statistically significant differences between the groups. Age was significantly associated with preferences in DDN in Greek monolinguals and immigrants when analysed separately: being older Greek monolingual or immigrant increases the chance of selecting the object instead of the subject antecedent in DDN. A similar but weaker association was also found between Age and the groups of HS and L2ers when considered together as one group.

In resolving NS in DIN, all groups again performed similarly at chance level with no statistically significant differences between the groups. Age was significantly associated with AP in DIN in Greek monolinguals and immigrants when analysed separately, as well as -less strongly so- in the groups of HS and L2ers when analysed together as one group. The associations showed that being older increases the chance of selecting the object over the subject antecedent.

7.3.3.2 Resolution of overt subject pronouns

In the resolution of OSP in the DDO condition, Spanish and Greek monolinguals were significantly different, with the Spanish group performing at chance level and the Greek-speaking groups manifesting a clear bias towards the object antecedent. The age of the participants was not significantly associated with their preferences in DDO.

Similarly, in resolving OSP in DIO, Spanish and Greek monolinguals were again significantly different. The Spanish group did not exhibit strong preferences, while the Greek group displayed a strong bias towards the object. Age showed a significant association in Greek monolinguals, signifying that being older increases the chance of selecting the object over the subject antecedent. In the case of immigrants, Age also revealed a significant association indicating that being older increases the chance of selecting the subject antecedent, contrary to what was observed in Greek monolinguals. In all Greek-speaking groups, however, the object was significantly favoured as the preferred antecedent of OSP.

7.3.4 Age

In order to explore Age over Antecedent Preferences (AP), follow-up binomial logistic regressions were performed introducing Age as independent variable in each group of speakers. Whenever a significant association was observed between Age and a particular group, then the finding was further examined.

7.3.4.1 DDN

In DDN, significant associations between Age and AP were found only in Greek monolinguals and immigrants. In both groups, Age revealed a significant negative association with AP. This shows that being older by one year increases the chance of selecting the object instead of the subject antecedent (OR=0.958; OR=0.968) by 4.4% and 3.3% in Greek monolinguals and immigrants respectively (Table 7.21).

Table 7.21. Binomial logistic regressions: Age and Antecedent Preferences in DDN

	OR	SE	z	p	OR CI	
Spanish	0.982	0.012	-1.490	n.s.	0.959	1.006
Greek	0.958	0.013	-3.140	0.002	0.933	0.984
IMM	0.968	0.008	-3.770	<0.001	0.952	0.985
HS	0.981	0.015	-1.320	n.s.	0.952	1.010
L2	0.978	0.016	-1.400	n.s.	0.947	1.009

(OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Greek monolinguals and immigrants

Since Age revealed a significant association with AP in Greek monolinguals and immigrants, the data for this condition were further elaborated examining the Age factor in younger and older speakers of these two groups. The monolinguals and immigrants were divided into a Younger (N=34) and an Older group (N=20) (see Kaltsa et al. 2015). The age in the Younger group was 16-58 (mean: 36.8, SD: 9.03) and in the Older group it was 60-87 (mean: 77.3, SD: 7.48). The Younger group consisted of the younger Greek monolinguals (N=13) and the younger immigrants (N=21), while the Older group consisted of the older Greek monolinguals (N=7) and the older immigrants (N=13). The results per age group for DDN are shown in Figure 7.5 and Table 7.22.

DDN

	Younger	Older
■ Subject	59.56	26.25
■ Object	40.44	73.75

Figure 7.5. Preferences in DDN according to age group (Greek monolinguals & IMM)

Table 7.22. Preferences in DDN according to age group (Greek monolinguals & IMM)

	Younger	Older	Total
Subject	81	21	102
	59.56	*26.25*	*47.22*
Object	55	59	114
	40.44	*73.75*	*52.78*
Total	136	80	216
	100	*100*	*100*

Tests of proportions showed that the preference for the subject antecedent in the Younger group, as shown in Figure 7.5, was statistically significant (z=2.190, p=0.029). In contrast, the Older group displayed a significantly higher preference for a match for the object antecedent (z=-3.837, p<0.001). The between group comparison revealed a statistically significant difference [Pearson χ^2 (1, N=216)=22.423, p<0.001].

When only the younger Greek monolinguals' preferences were examined, the subject was preferred (58.33%), but with no statistical difference between subject and object preferences (z=1.139, p=0.254). The older Greek monolinguals, by contrast, showed a clear bias towards the object (71.88%) and the difference between subject and object AP reached statistical significance (z=-2.267, p=0.023).

A very similar pattern was revealed when only the younger immigrants' preferences were considered. Namely, they also showed a preference towards the subject antecedent (60.71%), but the difference between subject and object AP was not statistically significant (z=1.919, p=0.054). The older immigrants selected the object significantly more often (73.08%) than the subject antecedent (z=-3.022, p=0.003).

In short, a difference was found between younger and older Greek monolinguals and immigrants. When the two younger subgroups were analysed together, they displayed a significant bias towards the subject. When each subgroup of younger speakers (i.e. younger Greek monolinguals and younger immigrants) was analysed separately, the preference towards the subject did not reach statistical significance. In contrast, older Greek monolinguals and older immigrants, analysed together as the Older group as well as separately, consistently prefered the object antecedent.

HS and L2ers

Although there was no significant association between Age and AP in HS and L2ers when analysed separately in DDN (Table 7.21), when considered together as one group there was a marginally non-significant association in regressions (Table 7.8) as well as a marginally significant association in ANOVA. This suggested that the older speakers of these groups selected the object antecedent more often. The speakers were thus divided into a Younger (N=28) and an Older group (N=13). The Younger group consisted of speakers aged 17-59 (mean: 40.4, SD: 12.2), while the Older group encompassed speakers aged 60-87 (mean: 67.4, SD: 6.5). The preferences of the two age groups revealed a difference shown in Figure 7.6 and Table 7.23.

While the Younger group exhibited random preferences in the DDN condition, the Older group significantly favoured the object antecedent (z=-3.418, p<0.001). The difference between Younger and Older groups of HS and L2ers was statistically significant [Pearson χ^2 (1, N=164)=8.138, p=0.004].

DDN	Younger	Older
■ Subject	46.43	23.08
■ Object	53.57	76.92

Figure 7.6. Preferences in DDN according to age group (HS & L2)

Table 7.23. Preferences in DDN according to age group (HS & L2)

	Younger	Older	Total
Subject	52	12	64
	46.43	*23.08*	*39.02*
Object	60	40	100
	53.57	*76.92*	*60.98*
Total	112	52	164
	100	*100*	*100*

The group of Younger HS/L2ers was significantly different when compared to the group of Younger Greek monolinguals/immigrants in DDN [Pearson χ^2 (1, N=248)=4.258, p=0.039]. The former group exhibited random optionality, while the latter group showed a bias towards the subject antecedent. On the other hand, the group of Older Greek monolinguals/immigrants was not significantly different when compared to the group of Older HS/L2ers in DDN [Pearson χ^2 (1, N=132)=0.169, p=0.681].

7.3.4.2 DIN

When the DIN condition was analysed with regard to Age, similarly to DDN, only Greek monolinguals and immigrants displayed significant associations between Age and AP. Age showed a significant negative association with AP, indicating that being older by one year increases the chance of selecting the object instead of the subject antecedent in the groups of Greek monolinguals (OR=0.975) and immigrants (OR=0.972) by 2.6% and 2.9% respectively (Table 7.24).

Table 7.24. Binomial logistic regressions: Age and Antecedent Preferences in DIN

	OR	SE	z	p	OR CI	
Spanish	1.015	0.012	1.200	n.s.	0.991	1.039
Greek	0.975	0.012	-2.030	0.042	0.952	0.999
IMM	0.972	0.008	-3.430	0.001	0.956	0.988
HS	0.976	0.015	-1.630	n.s.	0.948	1.005
L2	0.973	0.016	-1.710	n.s.	0.943	1.004

(OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval, LB=Lower Bound, UB=Upper Bound)

Greek monolinguals and immigrants

Since Age revealed a significant association with AP in Greek monolinguals and immigrants, as in DDN, the data were examined for the Younger and Older speakers of these groups. The results according to Age are shown in Figure 7.7 and Table 7.25.
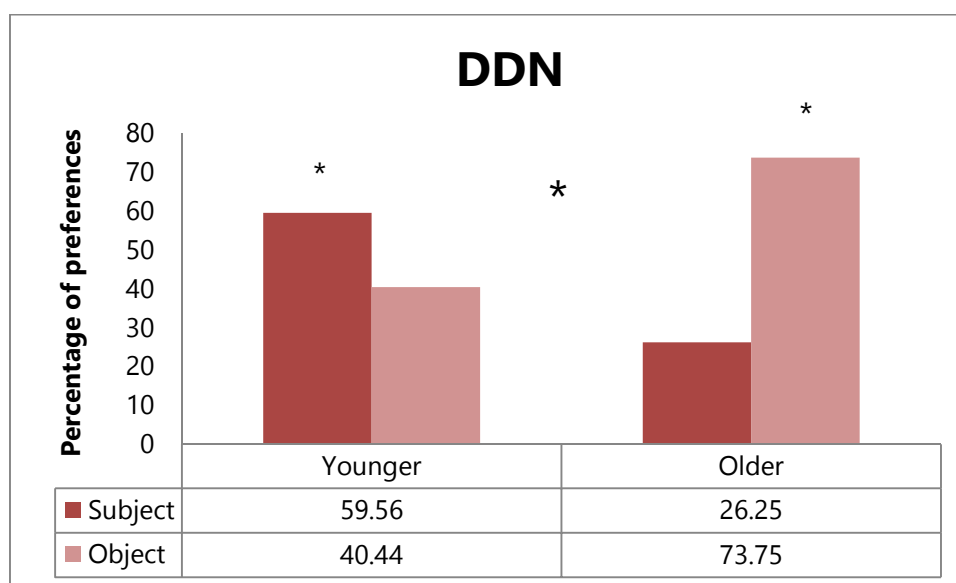

Figure 7.7. Preferences in DIN according to age group (Greek monolinguals & IMM)

Table 7.25. Preferences in DIN according to age group (Greek monolinguals & IMM)

	Younger	Older	Total
Subject	80	25	105
	58.82	*31.25*	*48.61*
Object	56	55	111
	41.18	*68.75*	*51.39*
Total	136	80	216
	100	*100*	*100*

Similarly to DDN, in DIN the Younger group significantly preferred the subject antecedent (58.82%) (z=2.026, p=0.043), whereas the Older group significantly favoured the object antecedent (68.75%) (z=-3.141, p=0.002). The comparison between the two age groups showed a statistically significant difference [Pearson χ^2 (1, N=216)=15.330, p<0.001].

When only the preferences of the younger Greek monolinguals were examined, again the subject was favoured (with the same percentage as in DDN, i.e. 58.33%), but the difference between subject and object AP was not statistically significant (z=1.139,

p=0.254). The older Greek monolinguals revealed a tendency to select the object antecedent (62.5%) but, contrary to DDN, the difference between subject and object AP was not statistically significant (z=-1.372, p=0.170).

When the responses of the younger immigrants were examined, the preference for the subject (60.71%) was not statistically significant (z=1.919, p=0.054) relative to the preference for the object. In contrast, the preferred interpretation of the older immigrants was towards the object (73.08%) and the difference between subject and object AP in this subgroup was statistically significant (z=-3.022, p=0.003).

In sum, a significant difference was observed between younger and older groups of Greek monolinguals and immigrants. The younger group attached the NS to the subject significantly more often than to the object. When each subgroup of younger speakers was analysed separately, in each case their preference for the subject was not statistically significant. On the other hand, when taken together as one group, the older Greek monolinguals/immigrants preferentially matched the NS to the object reaching statistical significance. When analysed separately, the older Greek monolinguals' preference for the object did not reach statistical significance, whereas the older immigrants exhibited a significantly strong preference for the object.

The participants' responses as per age group did not differ between DDN and DIN conditions [Younger group: Pearson χ^2 (1, N=272)=0.015, p=0.902; Older group: Pearson χ^2 (1, N=160)=0.488, p=0.485].

HS and L2ers

Although there was no significant association between Age and AP in HS and L2ers when analysed separately in DIN (Table 7.24), when considered together as one group (Table 7.21) a significant association was found: the older speakers tended to prefer the object antecedent. The results of the AR task were examined by dividing HS and L2ers into a Younger (N=28) and an Older group (N=13), in the same fashion as in the previous section. The preferences of the two age groups manifested differences shown in Figure 7.8 and Table 7.26.
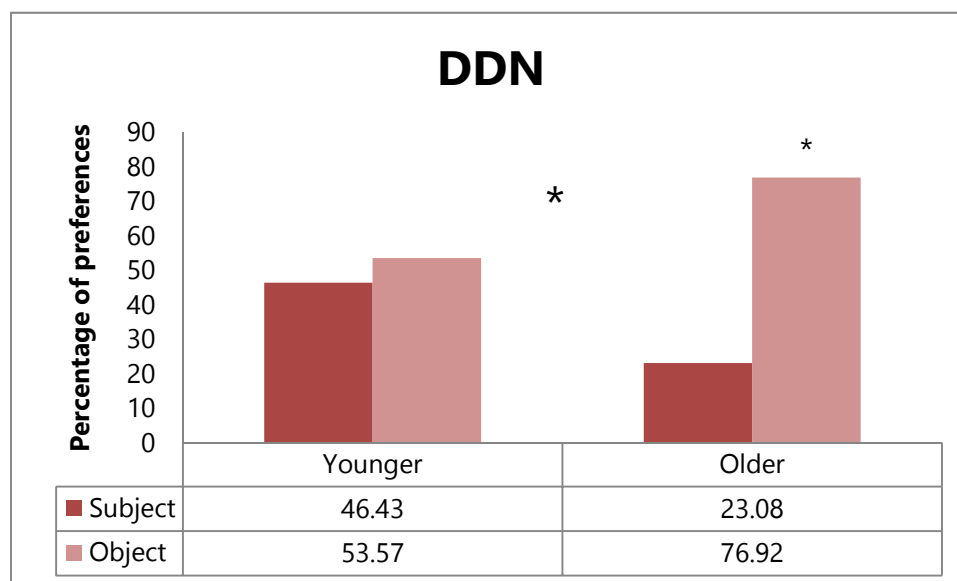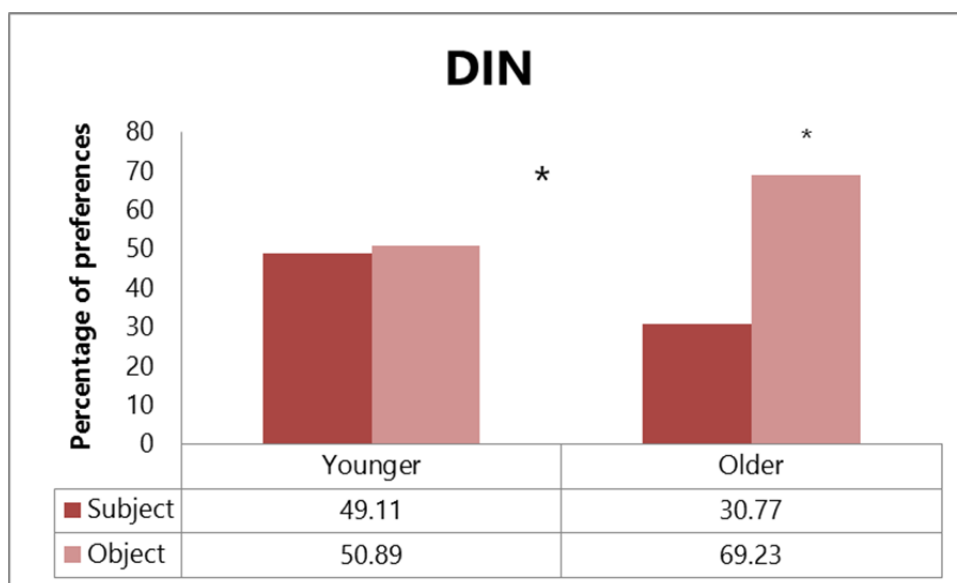
Figure 7.8. Preferences in DIN according to age group (HS & L2)

Table 7.26. Preferences in DIN according to age group (HS & L2)

	Younger	Older	Total
Subject	55	16	71
	49.11	*30.77*	*43.29*
Object	57	36	93
	50.89	*69.23*	*56.71*
Total	112	52	164
	100	*100*	*100*

As in DDN, in DIN the Younger group consisting of HS and L2ers did not have a bias towards an antecedent, whereas the Older group significantly favoured the object antecedent (z=-2.589, p=0.009). The difference between Younger and Older groups of HS/L2ers was statistically significant [Pearson χ^2 (1, N=164)=4.864, p=0.027].

The group of Younger HS/L2ers was not significantly different when compared to the group of Younger Greek monolinguals/immigrants in DIN [Pearson χ^2 (1, N=248)=2.338, p=0.126]. In addition, the group of Older monolinguals/immigrants was not significantly different when compared to the group of Older HS/L2ers in DIN [Pearson χ^2 (1, N=132)=0.003, p=0.953].

Interim summary

With respect to the NS resolution, Age played a significant role in the groups of Greek monolinguals and immigrants. While the younger speakers showed a clear bias towards the subject antecedent, the older speakers had the opposite behaviour consistently attaching the NS to the object antecedent. In the conditions with the NS, the HS and L2ers when grouped together also revealed considerable differences between younger and older speakers, with the latter preferring the object antecedent and the former performing at chance. The (in)definiteness of the matrix object did not affect the participants' responses. Thus, grouping the speakers according to two age groups (Younger, Older) uncovered that age has a significant impact on the resolution of NS.

7.3.4.3 DDO

In the analysis of the DDO condition, no significant associations between Age and AP were found (Table 7.27).

Table 7.27. Binomial logistic regressions: Age and Antecedent Preferences in DDO

	OR	SE	z	p	OR CI	
Spanish	0.995	0.012	-0.460	n.s.	0.972	1.018
Greek	0.971	0.016	-1.780	n.s.	0.940	1.003
IMM	0.998	0.011	-0.160	n.s.	0.977	1.020
HS	1.005	0.021	0.220	n.s.	0.965	1.046
L2	0.989	0.019	-0.590	n.s.	0.953	1.026

(OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval)

Greek monolinguals and immigrants

For comparison purposes, Figure 7.9 and Table 7.28 present the preferences of the Younger and Older groups consisting of Greek monolinguals and immigrants. The object antecedent was strongly preferred by the Younger (z=-6.4362, p<0.001) and the Older group (z=-5.1292, p<0.001) with no statistical difference between the two age subgroups [Pearson χ^2 (1, N=216)=0.135, p=0.713].
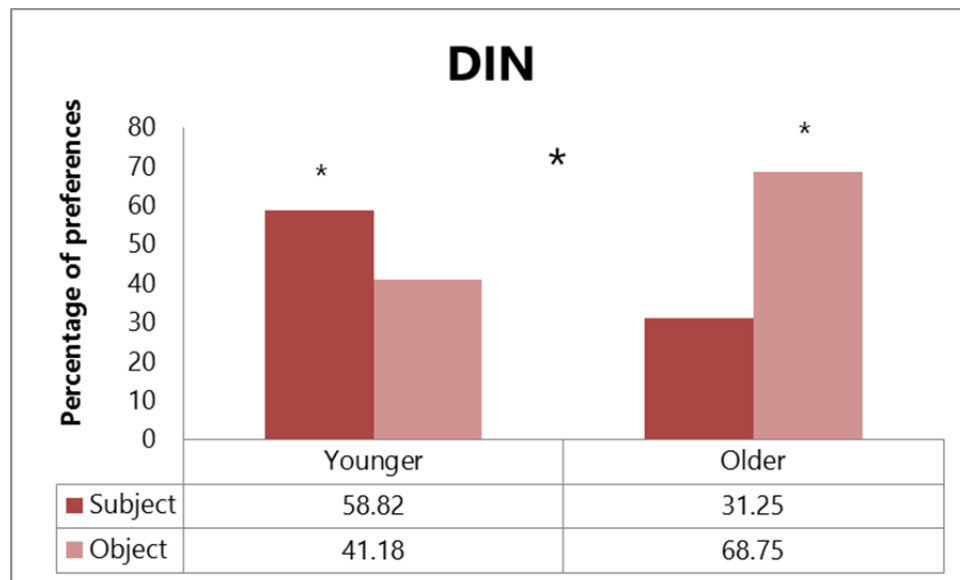
Figure 7.9. Preferences in DDO according to age group (Greek monolinguals & IMM)

Table 7.28. Preferences in DDO according to age group (Greek monolinguals & IMM)

	Younger	Older	Total
Subject	23	12	35
	16.91	*15*	*16.2*
Object	113	68	181
	83.09	*85*	*83.8*
Total	136	80	216
	100	*100*	*100*

7.3.4.4 DIO

In the analysis of the DIO condition, there were significant associations between Age and AP only in Greek monolinguals and immigrants (Table 7.29). In Greek monolinguals, Age showed a significant *negative* association with AP, indicating that being older by one year increases the chance of selecting the object instead of the subject antecedent (OR=0.955) by 4.7%. In the case of immigrants, Age showed a significant *positive* association with AP, indicating that being older by one year increases the chance of selecting the subject instead of the object antecedent (OR=1.033) by 3.3%.

267

Table 7.29. Binomial logistic regressions: Age and Antecedent Preferences in DIO

	OR	SE	z	p	OR CI	
Spanish	0.984	0.012	-1.30	n.s.	0.960	1.008
Greek	0.955	0.016	-2.75	0.006	0.924	0.987
IMM	1.033	0.012	2.74	0.006	1.009	1.058
HS	1.003	0.017	0.16	n.s.	0.971	1.036
L2	0.960	0.021	-1.85	n.s.	0.919	1.002

(OR=Odds Ratio, SE=Standard Error, OR CI=Odds Ratio Confidence Interval)

Greek monolinguals and immigrants

Greek monolinguals and immigrants were analysed being grouped into a Younger and an Older group, as in the conditions previously examined. The results according to age are shown in Figure 7.10 and Table 7.30.

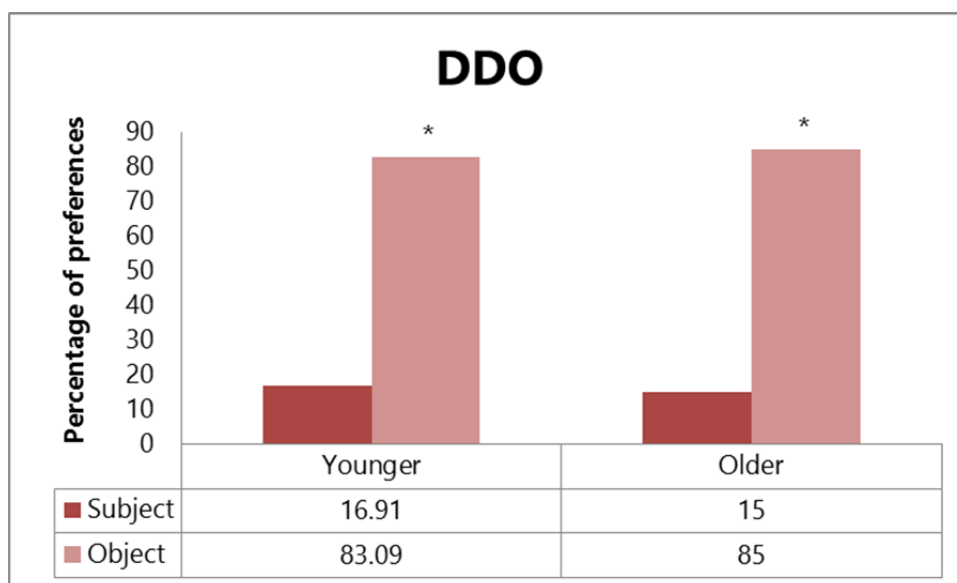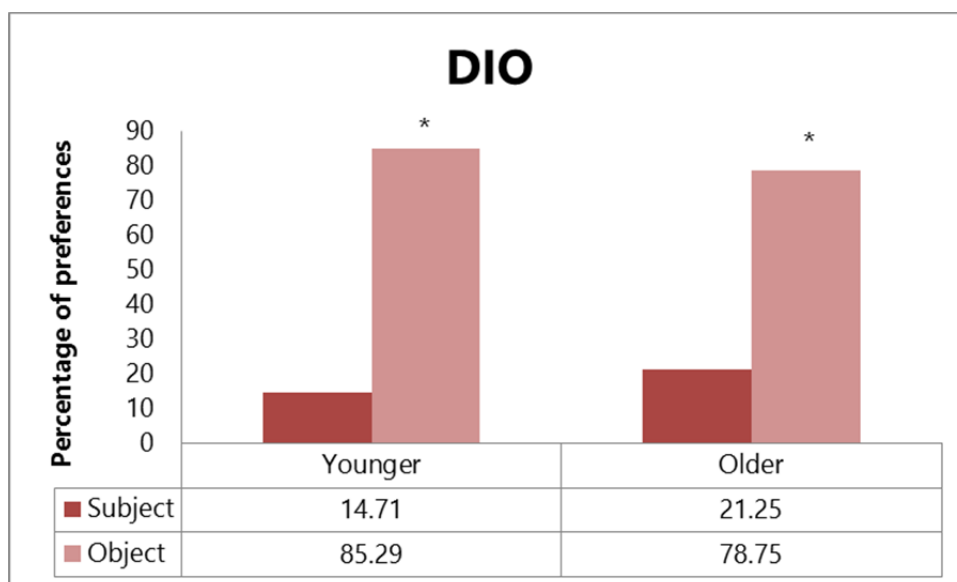Figure 7.10. Preferences in DIO according to age group (Greek monolinguals & IMM)

Table 7.30. Preferences in DIO according to age group (Greek monolinguals & IMM)

	Younger	Older	Total
Subject	20	17	37
	14.71	*21.25*	*17.13*
Object	116	63	179
	85.29	*78.75*	*82.87*
Total	136	80	216
	100	*100*	*100*

Both age groups strongly preferred the object antecedent (Younger: 85.29%, Older: 78.75%). This preference was statistically significant in both the Younger (z=-6.724, p<0.001) and the Older group (z=-4.458, p<0.001). There was no statistically significant difference between the groups [Pearson χ^2 (1, N=126)=1.519, p=0.218].

When only Greek monolinguals were considered, the object was consistently chosen by both younger (70.83%) and older speakers (87.5%) with no statistical difference between the groups [Pearson χ^2 (1, N=80)=3.059, p=0.080]. The same preferred interpretation of OSP emerged in the younger (92.86%) and older (75%) immigrants. However, the older immigrants chose the subject antecedent significantly more frequently than their younger counterparts [Pearson χ^2 (1, N=136)=8.52, p=0.004], hence the statistically significant positive association observed between Age and subject AP in the logistic regression. This behaviour reflects the choice of the first-mentioned referent as a plausible resolution strategy when the object antecedent was indefinite (less prominent). The difference between DDO and DIO was not statistically significant neither in the Younger [Pearson χ^2 (1, N=272)=0.249, p=0.618] nor in the Older group [Pearson χ^2 (1, N=160)=1.053, p=0.305].

In sum, in DIO the object was consistently preferred over the subject antecedent by the two age groups of Greek monolinguals/immigrants. Although the object was strongly preferred, in immigrants there was a significant difference between the preferences of the younger and the older speakers: both subgroups favoured the object, but the younger group did so significantly more than its older counterpart.

Interim summary

In resolving OSP, all Greek-speaking groups showed a strong preference towards the object antecedent, while the Spanish monolinguals manifested random optionality in both DDO and DIO conditions. Age effects were observed only in the DIO condition in monolingual Greek and immigrants. While the object was significantly favoured in DIO by both younger and older Greek monolinguals and immigrants, the older immigrants selected the subject antecedent significantly more often than the younger immigrants.

7.4 Demonstrative pronoun resolution in Spanish

There was a clear difference in the resolution patterns of the OSP between Greek and Spanish monolingual groups. This finding brought into focus the fact that the Greek third-person OSP is identical in form with the demonstrative, with ambiguous use between deictic and pronominal interpretation, whereas in Spanish this is not the case (see §2.3.6). The inherently deictic nature of *aftos* may thus render its use comparable to the use of the demonstrative *este* in Spanish. Along these lines, a follow-up AR study was conducted using an online task with Chilean Spanish participants only (N=20, age range: 27-85[17], mean: 45.1, SD: 16.1). The design of the task was the same as the one employed in the primary AR study with two differences:

(a) The experimental items were only those involving DDO and DIO contexts, but instead of the personal pronoun *él/ella* the demonstrative pronoun *este/esta* was included in the sentences, as in (93).

(93)   *La abuela besaba a la enfermera cuando <u>esta</u> se ponía el abrigo.*

   'The old lady was kissing the nurse when <u>she</u> was putting on her coat.'

(b) The task was designed as an online questionnaire, which was sent via a hyperlink to the participants by email. The test sentences were 8 (4 in DDO, 4 in DIO) and the ratio to the fillers was 1:1 (see Appendix D.5). The same male voice presented the sentences followed by the comprehension questions in the same fashion as in the main AR study. The participants were instructed to listen to the sentences only once and to respond as soon as they heard the question by clicking one of two given options involving the subject or the object antecedent. Nonetheless, differently from the main AR study, the participants could hear the sentences more than once since this was beyond control; hence, their responses could reflect metalinguistic processes. They also had to read and choose one of the two possible answers, which was not the case in the main AR study. Although the data elicitation method was different from that of the primary study, the results indicated that the demonstrative *este* in Spanish behaved like *aftos* in Greek, as shown in Figure 7.11 and Table 7.31.

---

[17] Binomial logistic regressions showed no significant association between Age and AP in DDO (N of observations: 80, z=0.66, p=0.506) and DIO (N of observations: 80, z=0.04, p=0.964).
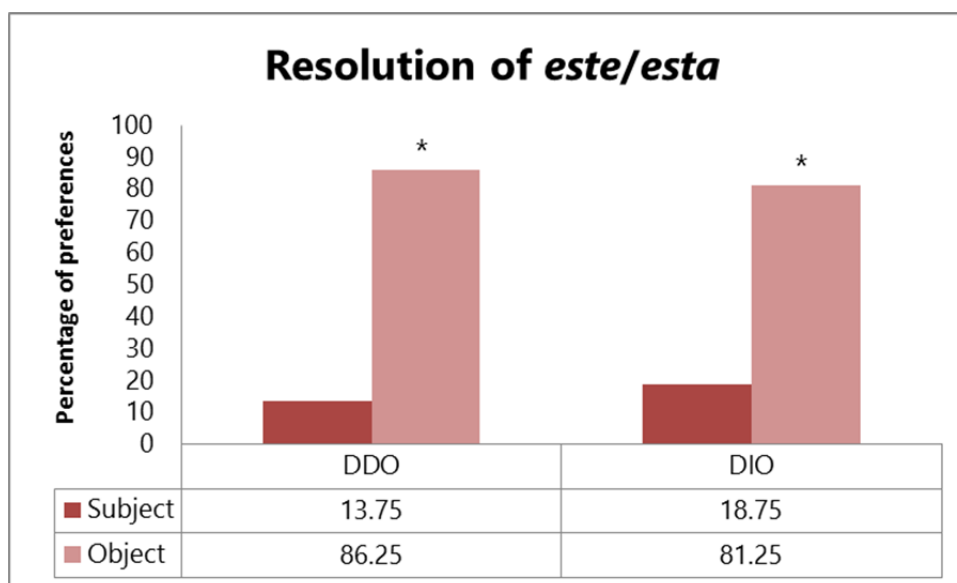
Figure 7.11. Spanish monolinguals: resolution of the demonstrative in DDO & DIO

Table 7.31. Spanish monolinguals: resolution of the demonstrative in DDO & DIO

	DDO	DIO
Subject	11	15
	13.75	*18.75*
Object	69	65
	86.25	*81.25*
Total	80	80
	100	*100*

The difference between DDO and DIO involving the demonstrative was not statistically significant [Pearson χ^2 (1, N=160)=0.735, p=0.391]. However, there was a statistically significant difference between resolutions of the demonstrative and the personal pronoun in Spanish in DDO [Pearson χ^2 (1, N=160)=25.641, p<0.001] and DIO [Pearson χ^2 (1, N=160)=9.643, p=0.002]. Moreover, there was no statistically significant difference between monolingual Greek in DDO and the respective Spanish responses involving the demonstrative [Pearson χ^2 (1, N=160)=0.427, p=0.514]. Similarly, no statistically significant difference was observed between Greek monolinguals in DIO and Spanish monolinguals in DIO containing the demonstrative [Pearson χ^2 (1, N=160)=0.344, p=0.558]. Thus, the Greek OSP *aftos* behaved like the Spanish demonstrative *este*, while both forms were at odds with the Spanish personal pronoun (*él/ella*), which triggered similar preferences to the NS in interpretation.

7.5 Summary of findings

7.5.1 Greek and Spanish monolinguals

In resolving non-biased ambiguous forward anaphora, Greek and Spanish monolingual speakers performed similarly in the conditions with NS (DDN, DIN) manifesting indeterminacy in their matching decisions. This suggests that the NS is ambiguous and does not always/necessarily signify TC. It was found, however, that the age of speakers played a significant role in resolving the NS in DDN, since older Greek monolinguals showed a clear preference towards the object antecedent, which was not the case for their younger counterparts. The age effect was not observed in the Spanish group. Definiteness of the matrix object did not seem to influence the resolution of NS in either group.

On the other hand, in OSP resolution (DDO, DIO), Greek and Spanish monolinguals behaved significantly dissimilarly. The Spanish speakers performed at chance level, similarly to NS resolution, whereas the Greek monolinguals strongly favoured the object antecedent. The age effect emerged only in DIO in the Greek monolinguals indicating that being older is associated with choosing the object antecedent more often. In sum, the OSP was clearly related to TS contexts in Greek and, by contrast, appeared to be ambiguous in Spanish. Definiteness of the matrix object apparently had no effect on the resolution preferences of OSP in either group. A follow-up study showed that the Greek OSP behaved like the Spanish demonstrative pronoun *este* in strongly establishing coreference with the object antecedent.

According to the findings of this study, Greek and Spanish did not perform along the lines of PAH on resolution of NS, since neither group manifested strong preferences towards an antecedent. Age emerged as an important factor in affecting resolution preferences of the NS in Greek since older Greek monolinguals significantly preferred the object referent in DDN. As for the resolution patterns of OSP, Greek monolinguals strongly preferred a disjoint interpretation in agreement with PAH. The Spanish-speaking group did not show clear resolution preferences for the personal pronoun, but their responses in resolving the demonstrative pronoun was in line with PAH.

7.5.2 Greek-Spanish bilinguals

Respecting NS (DDN, DIN) in non-biased ambiguous forward anaphora in Greek, the age played a significant role. It was found that younger immigrants patterned like younger monolingual Greek speakers performing according to the PAH, i.e. attaching the NS to the most prominent antecedent (the subject). By contrast, older immigrants similarly to older Greek monolinguals interpreted the NS as referring to the less prominent antecedent (the object). The younger HS/L2ers manifested a weakened PAH strategy in selecting both antecedents at similar rates, while their older counterparts preferentially selected the object, performing like the older immigrant/monolingual Greek speakers. (In)definiteness of the object did not affect the NS resolution preferences in the bilingual groups.

In resolving OSP (DDO, DIO), the bilingual groups performed comparably showing a strong preference towards the object antecedent, similarly to Greek monolinguals. Age effects were observed only in the DIO condition in immigrants. The older immigrants assigned the OSP to the subject antecedent significantly more often than the younger immigrants, unlike older Greek monolinguals. Nonetheless, the object antecedent was consistently preferred by all bilingual groups. Definiteness of the matrix object did not influence the resolution decisions regarding OSP in the bilingual groups.

According to the findings of this study, in NS resolution the groups of immigrants and L2ers generally manifested an overall performance which was similar to Greek monolinguals, i.e. NS showing no clear bias towards either antecedent. The group of HS in resolving NS manifested a more pronounced preference towards the object antecedent. Age effects emerged in the resolution of NS with older bilingual speakers preferring a recency-of-mention strategy for NS interpretation, similarly to older Greek monolinguals. The resolution patterns for OSP showed that the three bilingual groups performed comparably to Greek monolinguals and in line with the PAH, linking the OSP to the object antecedent most of the times.

7.6 Discussion

7.6.1 Comparison between Greek and Spanish

The first research question sought to answer whether Greek is different from Spanish in the distribution of third-person subjects in non-focused contexts. The hypothesis was that the two languages would differ in the interpretation of OSP. The Spanish OSP was predicted to be more variable than the Greek OSP because of the deictic nature of the latter, which renders it less flexible and less ambiguous by comparison.

7.6.1.1 Null subjects

The production data indicated that the OSP was mainly used in TS, while the NS was used in both TC and TS, although predominantly so in TC establishing coreference with the subject of the previous clause, according to the predictions. It could be reasonably expected that in AR the NS would mostly pick the subject antecedent, as also attested in several previous studies (see Chapter 4). However, Greek and Spanish monolinguals manifested unclear resolution patterns for NS on the whole (except the small group of older Greek monolinguals, who showed a significant preference for the object antecedent in DDN).

The indeterminate interpretation preferences in both Greek and Spanish suggest that the NS is ambiguous, variable and flexible exhibiting a relatively wide scope in both languages. Although morphologically underspecified, hence inherently ambiguous, NS can be felicitously used in TS when contextual and/or morphological cues allow reference disambiguation (see §2.4.4, §3.8). Nonetheless, in decontextualised sentences the prediction of the PAH proposes reliance on a default strategy, with NS interpreted as corefering to the matrix subject more often than not. This was not found in the present study in neither Greek nor Spanish, against previous research findings (Greek: Prentza & Tsimpli 2012; Kaltsa et al. 2015; Papadopoulou et al. 2015; Spanish: Alonso-Ovalle et al. 2002; Iverson 2012; Filiaci et al. 2013; Clements & Domínguez 2016), but in line with others (Greek: Dimitriadis 1996; Mastropavlou et al. 2014; Spanish: Callahan et al. 2007; Chamorro et al. 2015).

Both production and interpretation findings showed that NS do not always entail TC. The question, however, is why the AR results of this study differ from others. This could be attributed to the method in which the task was presented to the participants. For reasons explained in §5.3.4.2, the AR task was exclusively auditory, i.e. the participants heard the experimental sentences and replied orally without having access to any visual cues (e.g. text or pictures). It could be thus assumed that memory constraints possibly conduced to recency of mention playing a role, especially in the older speakers who took part in the study. Thus, participants may have chosen the object antecedent more often than expected because it was the antecedent most recently heard. However, Mastropavlou et al. (2014), i.e. the original study adapted for the present research, reported similar results on NS resolution using a picture-matching task. Namely, adult Greek monolinguals (aged 20-28) did not show a clear bias towards an antecedent in the NS condition. The current study, therefore, replicated the findings of the original study. In addition, Callahan et al. (2007) report a similar behaviour for NS resolution in speakers of Chilean Spanish, which is the particular Spanish variety under consideration here. Iverson (2012), however, did not find the same pattern in his Chilean Spanish control group, who preferred the NS referring to the preceding subject. The aforementioned Chilean-Spanish studies considered AR in different types of sentences (see §4.2.1.2).

7.6.1.2 Overt subject pronouns

In line with the production findings, the OSP in Greek was strongly interpreted as coreferring to non-prominent antecedents (i.e. objects), as has been consistently reported in the literature (Dimitriadis 1996; Miltsakaki 2007; Prentza & Tsimpli 2012; Mastropavlou et al. 2014; Kaltsa et al. 2015; Papadopoulou et al. 2015). On the other hand, although in Spanish the OSP was principally produced in TS contexts, the AR findings showed a variable behaviour in OSP interpretation, similar to that attested in NS. This accords with several Spanish studies which did not show clear resolution preferences for OSP (Alonso-Ovalle et al. 2002; Callahan et al. 2007; Keating et al. 2011; Jegerski et al. 2011; Filiaci 2011; Filiaci et al. 2013), but contrasts with others which report a bias in favour of the object antecedent (Gelormini-Lezama & Almor

2011; Iverson 2012; Shin & Cairns 2012; Chamorro et al. 2015). It is again noteworthy that the OSP findings of the present study corroborated earlier findings in Mastropavlou et al. (2014) for Greek and Callahan et al. (2007) for Chilean Spanish. Monolinguals of Greek and Spanish manifested significantly different resolution preferences of OSP, according to predictions. In line with production, differences in interpretation of OSP between the two languages were found in the context of TS.

In AR, while NS did not favour subject coreference to a considerable degree, thereby revealing variable behaviour, the predictions were confirmed as to the fact that more similarities were observed with NS than with OSP resolution between Greek and Spanish. Resolving OSP was significantly different between the two languages by virtue of the strong deictic feature of the Greek OSP. Filiaci (2010, 2011) found a similar divergence between Italian and Spanish OSP. While the Italian OSP (*lui/lei*) was consistently associated with TS, the Spanish OSP (*él/ella*) was easily processed in both TC and TS discourse contexts. She attributed this difference to the nature of pronominal forms. The Spanish personal pronoun is 'structurally deficient' or weak (Cardinaletti & Starke 1999), thus it may corefer to topic/subject antecedents (Filiaci 2010: 180). In contrast, the Italian OSP, like the Greek OSP, is a strong pronominal, hence strongly biased toward non-prominent antecedents.

Although the Spanish monolinguals' responses revealed chance performance as opposed to PAH in resolving the personal pronoun (*él/ella*), the demonstrative (*este*) was disjointly interpreted, consistently referring to the object antecedent, similarly to the Greek OSP *aftos*. Gundel (1996: 145) proposes that demonstratives indicate a referent which is 'activated (readily accessible to consciousness)', i.e. recently mentioned or available in the extralinguistic context. The recency constraint has been reported in the literature as an element which may determine reference resolution (e.g. Givón 1983; Sorace & Filiaci 2006; Prada Pérez 2009; Arnold et al. 2013). The Greek OSP *aftos* behaves like the demonstrative; hence its search space is more local, thereby identifying with the preceding object, similarly to the Spanish demonstrative pronoun.

In a nutshell, the OSP was clearly related to TS contexts in Greek and, by contrast, appeared to be unpredictable in Spanish. Definiteness of the matrix object had no effect on resolution preferences of OSP in either group, corroborating Mastropavlou et al.'s (2014: 26) claim that 'markedness' of OSP is 'stronger than definiteness of potential antecedents'.

Interim conclusion

Greek and Spanish monolinguals did not perform consistently with PAH on NS, since neither group manifested clear resolution preferences towards the subject referent. As for the OSP, Greek monolinguals strongly preferred a disjoint interpretation, in agreement with PAH. The Spanish-speaking group, on the other hand, interpreted the OSP more liberally, indicating that Spanish is significantly more permissive with establishing coreference between OSP and prominent antecedents (i.e. subjects) than Greek. This evidences the fact of the strong deictic nature of the Greek OSP *aftos*, which does not distinguish between personal and demonstrative forms (see §2.3.6.1).

7.6.1.3 The age factor

An increasing frequency of selecting the object antecedent along with age emerged in three out of the four conditions, namely in DDN, DIN and DIO, only in Greek monolinguals. An analogous finding is reported in Kaltsa et al. (2015), in which older and younger Greek monolinguals also exhibited significant differences in their NS resolution preferences. The younger group selected the subject significantly more often than the object, whereas the older group did not show a strong preference for an antecedent. The difference in findings between the present study and Kaltsa et al. (2015) probably stems from the different task design. Nonetheless, a similarity can be observed regarding the effect of age. In the present study, the younger speakers revealed an ambiguous behaviour while the older speakers preferably picked the object antecedent (in DDN). In Kaltsa et al. (2015), the younger speakers preferred the subject antecedent while the older speakers showed an ambiguous behaviour, i.e. they selected the object antecedent more often than their younger counterparts. In both studies, the older speakers manifested preferences which were inclined towards

the object antecedent with respect to the younger participants' preferences in resolving the NS.

Recency of mention is one of the factors which potentially make information salient (Arnold et al. 2013). Since AR taxes memory resources and older people generally have more limited memory capacity, non-linguistic aspects of cognition were evidently at play in tracking back antecedents. The older speakers' behaviour regarding NS resolution can be explained as recurring to recency as a convenient processing strategy. This is most likely related to the memory system, for which establishing short-distance referential dependencies is less taxing than doing so for long-distance dependencies. Assigning the NS to the most-recently heard antecedent can be thus regarded as a phenomenon associated with age-related weakening of cognitive processes, affecting working memory, since it was observed in the older Greek monolingual speakers.

The question arisen is why a similar age effect did not emerge in Spanish. Both NS and OSP forms in Spanish were flexible in AR allowing both subject and object referents as their antecedents and age-related cognitive ability did not seem to affect this process. Filiaci (2010, 2011) also found no difference in reading times for OSP in TC vs TS for Spanish (see also Alonso-Ovalle et al. 2002). In Greek, by contrast, NS and OSP have a clearer division of labour with the OSP favouring disjoint interpretations and the NS being more variable. Papadopoulou et al. (2015: 114) reported a subject bias for the NS, but considering performance times 'the resolution preferences of the null pronominal do not become apparent instantaneously' as opposed to OSP. NS in Greek allowed optionality and were susceptible to age-related memory and attention allocation constraints, which do not seem to be operative when there is no high competition between forms, as occurs in Spanish. Competing forms in a pronominal system (i.e. NS vs OSP in Greek) demand more processing resources in being distinguished and thus NS are subject to cognitive constraints, which may be related to age. Less competing forms in a pronominal system (i.e. NS vs OSP in Spanish) entail less cognitive effort in their distinction because both are acceptable in both contexts since neither of them imposes a strict bias; hence these forms are less susceptible to age effects. If the discourse/pragmatic distinguishing

property of the alternating options is not immediately established, both overt and null forms obtain random answers in interpretation of out of the blue contexts irrespective of participants' age/cognitive ability, as found in Spanish.

The above reasoning suggests that although resolution of ambiguous NS could be seen as a domain-general behaviour in older speakers, i.e. involving processing that is not language-specific, it seems to be also linked to aspects of language-internal processing depending on the particular language being examined. This question, however, remains open and needs to be addressed in further research, since the age effect in the Spanish group may have not emerged due to the group size.

7.6.1.4 Conclusion

Overall, the AR results were in line with previous research, confirming the predictions regarding OSP, but showing that NS were more variable than expected. The presumed TC contexts were found to be more 'ambiguous' than anticipated, i.e. interpreted as TS half the time in both languages. This corroborates the fact that NS are inherently ambiguous and variable since they can be accepted in both TC and TS. Differences between Greek and Spanish were found in the scope of OSP, being wider in Spanish and narrower in Greek. The Greek OSP was interpreted as marking TS and its behaviour was found to be similar to the Spanish demonstrative *este*. The age factor affected the Greek but not the Spanish monolingual performance, with recency of mention being a strategy to follow for the older speakers.

7.6.2 Comparison of Greek-Spanish bilinguals

The performance of bilinguals in AR in Greek was predicted to be more inconsistent than that of monolinguals. Resolution of OSP was thus expected to pick more often subjects, a prediction which was disconfirmed. Additionally, bilinguals' matching patterns in resolving NS were expected to be more unstable than in monolinguals. This was the case for younger HS and L2ers. Age emerged as a crucial variable in AR.

7.6.2.1 Null subjects and the age factor

In the bilingual performance, NS at first glance exhibited overall optionality in resolution preferences when no age distinction was made within the groups. This performance was similar to that of monolinguals, with the exception of HS who linked the NS to the object referent more often than the other groups. However, a distinction between age groups unveiled significant differences in resolution patterns between younger and older speakers. Greek monolinguals were considered together with immigrants and HS together with L2ers. This grouping ensued from patterns emerged in regressions conforming to basic sociolinguistic characteristics, such as being (more) dominant in Greek (monolinguals, immigrants) or speaking Greek as the weaker language of the pair (HS, L2).

The pattern found was the following: the older speakers, regardless of linguistic background, clearly preferred to match the NS to the object antecedent. The younger speakers did not show a homogeneous pattern. Their performance depended on their linguistic profiles, i.e. the younger attrited/monolingual Greek speakers tended to be consistent with the PAH, while the younger HS/L2ers exhibited inconsistent preferences. A weakened PAH strategy in bilinguals was one of the predictions, which was confirmed by younger HS/L2ers' performance in NS resolution. A weakened PAH strategy was also attested in the performance of all older participants. Since this was also found in older Greek monolinguals, it seems to derive from age and memory effects, i.e. from language-external cognitive mechanisms not directly related to bilingualism, pointing out the impact of age in ambiguous resolution decisions.

In Kaltsa et al. (2015), older monolinguals also patterned like attriters of similar age. Their preferences, however, favoured the subject antecedent. This difference in findings may be due to the different language pair and method. Additionally, in Kaltsa et al. (2015) HS' preference for the subject was significantly weaker than in monolingual controls. An analogous pattern emerged in the present study with younger HS/L2ers manifesting weak resolution preferences compared to their monolingual/immigrant counterparts.

The older participants, irrespective of being monolingual or bilingual, employed the strategy of selecting the object antecedent possibly because of age-related reduced cognitive efficiency of working memory and/or attention span, for which linear distance plays a pivotal role. As Hendriks et al. (2014) assert, elderly adults may be less responsive to discourse prominence of referents than younger adults and their performance may be more affected by intervening linguistic material between a subject pronoun and its antecedent. Antecedent preferences of older participants (monolinguals, immigrants, HS, L2ers) in NS interpretation in Greek point to recency of mention as guiding their resolution preferences. Kaltsa et al. (2015) also found that older attriters/monolinguals showed a significant tendency to pick the object as corefering to NS compared to younger HS/monolinguals in Greek. This may be a consequence of older speakers' inefficient inhibitory processes, which allow irrelevant information into working memory creating interference (Abrams & Farrell 2010).

For the younger immigrant/monolingual Greek speakers, resolution of NS was in line with the PAH, as in several previous studies on Greek (see §4.1). On the other hand, the younger HS/L2ers revealed an unclear pattern, similar to Spanish monolinguals: the NS appeared to be more ambiguous than in participants of similar age who were monolingually raised in Greek. Thus, due to language contact in conditions of reduced exposure to Greek, matching decisions were irresolute in speakers who interpreted ambiguous NS anaphora in their weaker language. In addition, it was observed that there was a difference between the mean ages of younger immigrants/monolinguals (36.8) and the younger HS/L2ers (40.4), with the latter being a few years older than the former (see §7.3.4.1). This fact could also explain to some extent why the younger HS/L2ers (older than their immigrant/monolingual counterparts) exhibited more preferences towards the object antecedent, which was found to be a characteristic of older speakers in general.

The fact that HS and L2ers patterned similarly evidences the fact that 'heritage speakers and L2 learners share the same dominance pattern regardless of in what sequence the languages were acquired' (Montrul 2016b: 250; see also Montrul 2008). Although HS were simultaneous/early bilinguals and L2ers were sequential/late bilinguals, in both cases Spanish was dominant and Greek was the weaker language.

Reduced input in Greek across the lifespan in a Spanish-dominant social environment along with weaker proficiency compared to monolingually raised speakers may account for differences in performance between these two larger groups of speakers. It should be reminded, however, that in production the older immigrants performed similarly to HS in using the NS in TS more often than the other groups, indicating speaker-internal constraints in considering listeners' perspective due to processing limitations, which results in an apparent less fine-grained sensitivity to discourse context. The L2ers, on the other hand, tended to be over-descriptive in production.

7.6.2.2 Overt subject pronouns

OSP was expected to be affected by language contact surfacing as overextension of its scope in interpretation in the bilingual performance. Specifically, it was predicted that bilinguals would attach the OSP to the subject antecedent more often than monolinguals. This prediction, however, was disconfirmed since the three bilingual groups robustly preferred attachment to the object antecedent, indistinguishably from Greek monolinguals and differently from Spanish monolinguals.

This finding renders support to the observations on the production data. In both production and AR, the OSP was resistant to crosslinguistic influence. Despite its interface properties, which render it vulnerable in bilingual domains, the Greek OSP in contact with Spanish did not allow more optionality than in the monolingual baseline, possibly due to its strong grammatical properties. This is against predictions of the IH, according to which language contact, even involving pairs of NS languages, would result in the OSP being more open to coreferential readings with either subjects or objects referents. This is also against the findings in Kaltsa et al. (2015), where the OSP was indeed affected in Greek in contact with Swedish (a non-NS language) as manifested in AR preferences of HS and attriters. In the case of Greek, the other language of the pair apparently plays a role, i.e. a non-NS language seems to trigger stronger effects than a NS language, such as Spanish. Therefore, overextension of the scope of OSP does not seem to occur 'regardless' of whether the bilinguals' other language is a NS language or not (Sorace 2011: 14), at least in the case of the Greek performance of bilinguals.

7.6.2.3 Conclusion

The findings of Study 2, together with those of Study 1, support the prediction of the representational account, according to which no differences between monolinguals and bilinguals are expected in the scope of third-person OSP in combinations of prototypical NS languages. No influence from one grammatical system to the other obtained because the similarity of the pragmatic settings in the languages involved did not allow underspecification of the interpretable features of [Topic Shift] and [Focus]. Although variation at a microparametric level was clearly observed between the two languages, both Greek and Spanish are consistent NS languages, thus in such cases it seems to be less likely that the relevant features become underspecified due to attrition. Consequently, the OSP in Greek remains unambiguous in both production and interpretation in the performance of bilinguals. This brings about the comparison between the present study and other studies showing vunerability of the OSP, attrition and the effects of language pairs, which is discussed in §8.4.

8. General Discussion

The present chapter further discusses the research findings bringing together the production and interpretation results. It is divided into six sections. Section 8.1 discusses the findings obtained from Greek and Spanish monolinguals. Section 8.2 discusses the findings obtained from the bilingual groups. Theoretical implications are also discussed in Section 8.3 and comparisons with relevant studies are offered in Section 8.4. The major findings and conclusions are summarised in Section 8.5. Finally, Section 8.6 considers both linguistic and sociolinguistic factors relevant to the findings of the present research.

8.1 Monolinguals: Synthesising production and interpretation findings

Production and interpretation data in the two languages showed that NS were mainly used in TC and less frequently (yet considerably) in TS, while in AR the acceptance of NS was equally divided between TS and TC contexts. This shows a discrepancy between the two modalities, indicating that reduced forms, i.e. NS, preferentially establish coreference with prominent antecedents (i.e. subjects) in production. By contrast, the PAH strategy for NS was weaker in ambiguous contexts. This difference was against the predictions of the present study since such strategies were expected to be weakened in production because of the presence of context, which reduces possibilities of ambiguity, and more operative in interpretation due to lack of context. The weak preferences for NS assignment to subject antecedent may arise from the task design since similar findings were reported in Mastropavlou et al. (2014). In any case, NS were ambiguous in AR but crucially they were not ambiguous in production, even in TS. This suggests that the context plays a substantial role in reference interpretation as attested in monolingual production of NS. Pinto (2014), who also found TS in combination with NS, observes that such instances occur in environments which supply contextual cues for the identification and interpretation of NS. Thus, the inherent ambiguity of NS ceases to exist within particular contexts which provide the appropriate information to make its use contextually felicitous. Felicitous NS in TS in

production indicate that NS can be also linked to prominent antecedents which are not necessarily preceding subjects (see Miltsakaki 2007; Frana 2017).

The Greek production and interpretation data were consistent on the scope of OSP, which was produced in TS contexts and disjointly interpreted in AR. The low percentages of responses assigning the OSP to objects in AR affirm that resolution decisions, even for the inherently marked Greek OSP, may show strong tendencies but are not categorical (e.g. Tsimpli 2011). This means that the Greek grammar allows optionality to some degree with OSP. Put differently, although the Greek OSP is deictically marked, it is not inherently marked for TS. That being said, OSP resolution preferences in Greek were much stronger than those of NS. This was also illustrated in production, where the use of OSP was constrained by specific discourse features (TS, focus) with no deviations, while NS were more variable.

On the other hand, a difference was detected in Spanish between production and interpretation of OSP. While it was always felicitously used in TS (and in TC conveying focus), in AR it was equally accepted in TC and TS contexts. The discourse behaviour of the Spanish OSP was thus similar to that of NS in that it may be ambiguous when decontextualised but normally unambiguous within particular contexts. The fact that in Spanish the OSP was used in TS even when its presence was not absolutely necessary for reference disambiguation also indicates that it does not obey exactly the same pragmatic restrictions as the Greek OSP; hence, the empirical orbit of the Spanish OSP seems to be somewhat wider than in Greek. The findings of the present study suggest differences in the division of labour of pronominal subjects between Greek and Spanish, which seem to be related to a different morphological make-up of the two languages.

As Sorace (2011: 25) argues, it seems appropriate to 'differentiate structures on a gradient according to [...] whether they are closer to the "strongly biasing" syntactic end or to the "weakly biasing" contextual end'. It is assumed that OSP in both Spanish and Greek need to satisfy both syntactic and pragmatic conditions. However, if we were to put Spanish and Greek OSP in a gradient, as proposed by Sorace (2011), then the Greek OSP would be located towards the 'strongly biasing syntactic end'

with respect to the Spanish OSP, which would be situated closer to the 'weakly biasing contextual end'. Although both forms depend on discourse pragmatic factors, hence interface conditions, this is more evident in the Spanish OSP, since it was found to be ambiguous in interpretation and relatively flexible in production compared to the Greek OSP. Thus, the distribution of the Spanish OSP seems to be determined by the context more actively than in the case of the Greek OSP, whose use instead is a more clear-cut phenomenon. Foreshadowing the discussion about the bilingual data, the Greek OSP, due to falling more on the grammar side of the spectrum, carries anaphoric properties which make it more stable, or rather invulnerable, in language contact situations involving another NS language.

In sum, TC was found to be uncomplicated in production, encoded in similar ways in both Greek and Spanish. NS presumably encoding TC were controversial in interpretation of ambiguous anaphora giving rising to inconsistent AR preferences in both languages, with recency of mention guiding older Greek speakers' resolution strategy. TS revealed significant differences between Greek and Spanish in production and interpretation of OSP. While in Greek the OSP was consistently used and interpreted as referring to non-prominent antecedents, in Spanish there was a notable discrepancy between OSP production, which was virtually exclusive to TS, and OSP interpretation, which was referentially undetermined between TS and TC. In the monolingual performance, the effect of age did not emerge in production to the same degree as in interpretation (cf. Abrams & Farrell 2010; Shafto & Tyler 2014).

8.2 Bilinguals: Synthesising production and interpretation findings

In production, NS were used by bilinguals in a pattern which was generally similar to monolinguals in both TS and TC contexts, but with older immigrants and HS overusing NS in TS. In interpretation, NS were more ambiguous since they could encode both contexts, with all older speakers and HS preferentially interpreting them as encoding TS. This was possibly due to linear ordering of referents and working memory constraints in resolving ambiguity based on auditory perception skills, as well as processing overload in having two active languages. The greater the distance

between the pronoun and its antecedent, the more working memory is required for successful pronoun resolution in interpretation.

Distance may also affect the appropriate selection of subject form in production (Shin & Cairns 2012). Ambiguity attested in bilingual performance in using NS in TS indicates a more relaxed distribution of NS in bilingual grammars probably due to production constraints at the level of language use (processing), which induce weak monitoring of speech. Specifically, reducing the load on the production system favours NS as the computationally less costly option. Thus, NS distribution and interpretation were generally derived from economy strategies guided by the need to alleviate processing burden as reflected in preferences for less demanding processing operations. Such effect emerging in older immigrants and HS could be seen as a sign of attrition resulting from long-term language contact and conditions of reduced input and disuse. The same effect was also evidenced in the case of older monolinguals, who significantly favoured use of NS in TC contexts, but without resulting in ambiguity, and interpreted NS as encoding TC in AR compared to younger monolinguals. These findings imply that reduced monitoring involving the preference for NS in production as well as the 'recency of mention' strategy in interpretation are not necessarily or exclusively related to bilingualism. That being said, bilingualism may enhance such linguistic behaviours. As opposed to NS, the OSP was much more stable and consistent in both production and interpretation in all speakers, preferentially used and interpreted as marking TS.

The loose scope of NS in bilinguals' production and interpretation could be accounted for as a phenomenon of simplification (e.g. Silva-Corvalán 1991). NS were referentially complex because of their morphological underspecification and dependence on context. NS are used for economy purposes being licensed and identified in TS if the referent is sufficiently salient. Ambiguous NS, however, are not identifiable because their referent is not accessible in the state of knowledge of the interlocutor. In production, this means that the factor of salience was overridden. Loss of the pragmatic constraint of salience that guides linguistic choice, i.e. disregarding the degree of entity's salience in the discourse, which may result in incoherence, can be seen as a form of simplification in the performance of some bilinguals. Such

simplification could be further extended to interpretation of NS in AR, in which the salience factor as instantiated by subjecthood in Greek (Miltsakaki 2001) is lost or weakened granting recency of mention a more relevant role. This phenomenon triggers optionality in bilinguals and in older speakers more generally in the resolution of NS leading to a tendency for disjoint interpretation.

The idea of simplification is only suggestive with respect to the present data. It cannot be determined whether this process results in permanent consequences in some bilinguals' oral discourse or it is a rather non-permanent phenomenon related to real-time processing and 'moment-by-moment fluctuations in attentional control within individual speakers' (Sorace 2016: 9). The AR task was crucially based on the auditory attention of the participants. Extending Sorace's (2016: 7) idea on monolinguals, bilinguals as well 'may be occasionally unable or unwilling to engage in full processing', especially when the context is unambiguous to them in production and/or when interpreting globally ambiguous contexts. This may be enhanced by bilingualism effects, thereby bilinguals underperforming in their perspective-taking ability more often than monolinguals and tending to choose the most economical strategies. It is, therefore, plausible that performance factors (allocation of attention resources) were responsible for the attested misuse of NS.

Crucially, NS were found to be pragmatically susceptible, but OSP remained intact in language contact with Spanish insofar as bilinguals did not use/interpret OSP more extensively than Greek monolinguals. In the bilingual production, the design of the structures did not involve OSP as the default and in interpretation the robust preferences of bilinguals were indistinguishable from those of monolinguals. This indicates that the Greek OSP carries disambiguating lexical features derived from its very strong deictic function/meaning, which are not weakened in contact with another NS language. Juxtaposition of Greek and Spanish OSP indicated that the Greek OSP as a discourse element was more grammatically determined due to its more constrained use in very specific contexts. It was used to encode TS only when it was essential to disambiguate non-salient referents or to convey focus. In contrast, the Spanish OSP was more pragmatically determined since it could be used and interpreted more liberally. However, the difference between the OSP in the two

languages was not radical, hence crosslinguistic influence from Spanish did not obtain. The present findings suggest that use and interpretation of the Greek OSP is a more consistent phenomenon than use and interpretation of NS. Since OSP use is regulated by more precise rules, its discourse behaviour is more clearly established. This fact seems to make it less vulnerable in situations of contact where differences with the other language on the scope of OSP, although existing, are relatively weak.

In Greek, differences between Greek-Spanish bilinguals and monolinguals reside at the level of discourse/pragmatics in the scope of NS. The differences relate to language use and are not due to, but rather enhanced by bilingualism, since the NS is allowed by the Greek grammar as an economical option. Discourse/pragmatic distribution of OSP was not a source of optionality in adult monolinguals and bilingual speakers. This is arguably attributed to the strong deictic nature of the Greek OSP, which makes it more categorical hence less complex and (more) invulnerable to cognitive constraints due to age effects and/or contact with Spanish.

Finally, as for the age effects, the asymmetric pattern, with older adults experiencing greater difficulties in production compared to comprehension (Abrams & Farrell 2010; Shafto & Tyler 2014) is partially supported by the fact that the older immigrants produced more cases of genuine ambiguity. However, interpretation was also affected by aging since all speakers of Greek were prone to prefer the object antecedent in resolving NS.

8.3 Theoretical implications

The findings of the present study show that NS, constituting cases of higher ambiguity, were more affected than OSP in the performance of bilinguals in both production and interpretation. Specifically, NS were overused by HS and older immigrants in TS contexts and were also misused to some extent by the three bilingual groups giving rise to ambiguity. Additionally, in interpretation NS were found to be more permeable to age effects, as found in the performance of all Greek-speaking groups, and possibly to language contact, as found in the performance of HS and L2ers. The latter groups being sensitive to input from attrited speakers is also

a possibility related to their unstable performance in interpreting NS. By contrast, OSP were felicitously produced in both TS and TC contexts with neither quantitative nor qualitative differences between the Greek-speaking groups. Moreover, OSP were felicitously interpreted as marking TS by all the Greek-speaking groups. This strongly suggests that the scope of the Greek OSP in contact with Spanish remained unaffected to language contact and/or crosslinguistic influence, while NS, being more variable in production and interpretation, were more prone to interface vulnerability.

Such findings are in striking contrast to IH, which would predict significant differences between monolinguals and bilinguals only regarding the OSP, while the NS would be the unaffected form (Sorace 2011, 2012). This is because OSP are considered to carry properties which require integration of linguistic and pragmatic/cognitive information and imply processing constraints because they are externally interface-conditioned, whereas the NS are not, thereby being assumed to involve simpler informational structure. Crucially, however, in this study the NS displayed significant differences between the performance of monolinguals and bilinguals while OSP did not. While the attested differences were most likely due to performance shaped by processing constraints, these constraints did not affect use and interpretation of OSP.

Clements and Domínguez (2016) argue that IH, among other related generative research, has neglected the fact that NS can be used in TS contexts if the referent is salient enough. The present study showed that NS were used in TS contexts by both monolinguals and bilinguals and this use was mostly felicitous. However, instances of NS in TS in the bilingual performance sometimes resulted in ambiguity. This means that NS exhibit referential complexity because they can be used in TS contexts, but they are underspecified for reference, hence determined by contextual properties and thus susceptible to misuse. Misuse of NS may compromise communication generating misunderstandings or ambiguity, as opposed to overuse of OSP, which being infelicitous does not obscure or considerably slow comprehension (see Rothman 2009a; Sorace et al. 2009; Arnold et al. 2009; Judy 2015; Lozano 2018). In any case, overuse of OSP was not attested in the present study. The expected overuse of overt forms in production was instantiated by more extended use of LS, which could be regarded as a default referential expression in narratives. Crucially, NS also

showed an unpredictable behaviour in interpretation, being the culprit of differences between monolinguals and bilinguals as well as between younger and older age groups, whereas the OSP was firmly established in all cases. The findings point to the fact that NS are complex, in line with Clements and Domínguez (2016), more vulnerable than OSP, thereby being the root of variation.

The present findings support the idea of 'giving up neat dichotomies such as "narrow syntax vs. interfaces" but also neat distinctions between syntax-pragmatics, syntax-semantics, etc. and focusing on computational complexity as a key factor' (Sorace 2012: 210). In particular, it seems plausible to argue that null and overt subject alternation in general is a phenomenon residing at the syntax-discourse/pragmatics interface since it is determined by discourse/contextual factors (see Montrul 2011), involving the three forms of encoding third-person referents (NS, LS, OSP). Additionally, at least in the Greek performance of monolinguals and Greek-Spanish bilinguals, NS were more variable, hence more inconsistent in the input and apparently entailing higher degrees of computational complexity than OSP. This could be explained by interface conditions if we assume that NS are also complex at the syntax-discourse/pragmatics interface (see Clements & Domínguez 2016).

An alternative theoretical view which may explain such findings is the Vulnerability Hypothesis (VH) put forward by Prada Pérez (2018). This recent hypothesis takes into account the non-categorical distribution of subjects in NS languages such as Spanish as attested in variationist analyses. Accordingly, there is a multitude of factors affecting subject distribution, among which discourse context is a pivotal variable. VH establishes a categorical-variable continuum of vulnerability to language contact effects, in which 'structures that show variable distributions are permeable while those that exhibit categorical distributions are not' (ibid: 1). The prediction of the VH is, thus, that bilinguals would be similar to monolinguals in the most categorical variables and would differ in the least categorical variables, determining that complexity/variability of linguistic distributions is a better predictor than interfaces.

The VH prediction applies to the findings of the present study since the most categorical variable was the OSP in Greek, whose distribution was similar in

monolinguals and bilinguals. Overall, the OSP was used in TS or focused contexts practically 100% of the time in production and selected as marking TS more than 75% of the time in interpretation. On the other hand, the most variable distribution was observed in NS in Greek, which were used in TC in more than 86% of the contexts but also in TS in at least 21% of the contexts. In AR, additionally, NS distribution was nearly 50% in each context. Thus, use of NS was a highly variable phenomenon contingent on the discourse context and triggering differences between bilinguals and monolinguals as opposed to OSP, thus shown to be more vulnerable than OSP.

The same predictions can be considered regarding the discourse context instead of the subject form in the production data and again the VH is relevant. Specifically, TC was less variable than TS since it was prominently expressed by NS and was the least problematic context since no ambiguity emerged in TC. On the other hand, TS was more variable since it could be encoded by three options, namely LS, NS and OSP, depending on the status of the discourse referent's salience. This is in accord with Shin's (2014: 305) assessment of relative complexity in grammatical patterns on the following basis: 'the more choices of variants, the more complex the pattern'. It was found that TS was more problematic in the performance of bilinguals since only in TS contexts emergence of ambiguity could blur the intended meaning. Therefore, the claim of the VH that 'categorical distributions are not affected by language contact, while variable distributions may be affected by language contact' (Prada Pérez 2018: 21) is relevant and may account for the findings.

In determining selectivity of vulnerable domains in bilingualism, constraints stemming from the syntax-discourse/pragmatics interface were indeed found to be particularly complex as argued by Sorace (2011, 2012). However, these constraints can be viewed as applying to subject distribution more generally in NS languages. This claim is supported by the finding that in both monolingual and bilingual grammars it was the NS, not the OSP, which was pragmatically complex in Greek since the former and not the latter could more easily appear in both TS and TC contexts. The present findings are inconsistent with Sorace's IH (2011, 2012) in its current version, supporting the view of Montrul (2011: 603) that 'what we currently define as complex or difficult might not only be exclusively explained as a property of

a particular linguistic interface, but may as well be due to difficulty at many levels: processing and task demand, input factors and linguistic complexity'. The findings lend support to Prada Pérez's (2018) VH, whose predictions on language contact effects draw on the relative frequency of linguistic distributions, which establishes the variability, hence complexity, of the linguistic phenomena.

If we quit thinking in monolithic dichotomies (Sorace 2011, 2012; White 2011), it can be seen that certain phenomena may be variable and complex, thus vulnerable in language contact situations, because of syntax-discourse/pragmatics interface conditions. In other words, in order to understand particular linguistic behaviours, interface accounts along with vulnerability-according-to-variability accounts are not mutually exclusive. They are rather complementary since a complex interplay between the two seems to determine certain linguistic phenomena, such as subject distribution in monolingual and bilingual Greek.

With both linguistic and processing factors operating in tandem in production and interpretation of subject reference, the discourse pragmatic distribution of subjects was found to be indeed a source of optionality in adult speakers. In contrast to other studies, the most affected form in Greek, as shown in the present findings, was the NS, this being related to factors of computational efficiency. The OSP remained uninfluenced by the vulnerability caused by interface conditions in language contact. The invulnerability of the OSP can be also explained by the representational account (see §3.2, §6.6.2.1, §7.6.2.2), since no weakening of pragmatic constraints (discourse-interpretable features) occurred affecting the scope of the Greek OSP. According to this account, no L1 attrition is predicted in cases of language combinations of consistent NS languages, which instantiate a similarly complex pragmatic setting for the regulation of NS and OSP (see §8.4). Therefore, the postulations of the representational account are supported by the findings of this study as regards the non-existence of attrition signs, particularly in the linguistic behaviour of first-generation immigrants and HS, in the scope of the third-person OSP.

8.4 Comparison with other studies

In the context of L1 attrition, previous studies involving Greek have shown that the third-person OSP was indeed affected by language contact when the other language was a non-NS language (Tsimpli et al. 2004; Tsimpli 2007; Kaltsa et al. 2015). The fact that the present study, involving two NS languages, did not reveal any signs of attrition in the performance of bilinguals with respect to the scope of OSP suggests that the language combination matters and brings to the surface the issue of microparametric variation. This means that the language pair plays a key role, but any combination of NS languages would not necessarily give rise to the same effects. The language-specific nature of OSP as well as microparametric variation at other levels may also play a non-trivial part. In this regard, particular attention should be paid to the given language of the pair under examination, e.g. Greek, Spanish or Italian, bearing in mind that there may be crosslinguistic differences in the languages with respect to the same phenomenon, such as subject distribution (see Roussou & Tsimpli 2006). Furthermore, the particular domain of bilingualism is also important, i.e. whether the performance of speakers involves a developmental stage of L1/L2/simultaneous bilingual acquisition or L1/L2/heritage language attrition.

Bilingual contexts where OSP overuse/overacceptance was reported involving Greek and Spanish (Margaza & Bel 2006; Lozano 2006a, 2018; Georgopoulos 2017) concerned L1 Greek - L2 Spanish focusing on learners' L2 performance (see §3.4.1.2). It was therein documented that overgeneralisation of OSP in L2 Spanish tends to evaporate along with reaching higher proficiency levels, i.e. it is a developmental issue, hence only a passing phase. The present study had a distinct focus in examining different Greek-Spanish bilingual domains involving end-state grammars and concentrating on speakers' performance in Greek. It was also different from Sorace et al.'s (2009) study on bilingual children in that, aside from the different language combination and methodological design, it concerns adult speakers. That being said, the last-mentioned study yielded interesting findings concerning a different pair of prototypical NS languages and thus should be further commented.

The crucial finding in Sorace et al. (2009) is that the Italian-Spanish bilingual children overaccepted OSP in TC, which can be probably explained by the fact that 'the scope of the overt pronoun in Spanish is actually wider than in Italian' (ibid: 474). This microparametric difference was actually corroborated by the results in Filiaci (2011) and Filiaci et al. (2013) reporting differences in resolution of OSP in Italian and Spanish. It should be also noted that microparametric variation has been shown to obtain at other levels of grammar across Italian and Spanish. In Roussou and Tsimpli (2006), it is demonstrated that the VSO order is allowed in Spanish in cases of topicalisation or focus in clause-initial position (in Spec TP) (see Zubizarreta 1998; Zagona 2002). By contrast, Italian does not readily allow for VSO, since this word order is possible in more restricted cases, e.g. when the subject bears contrastive focus (see Belleti 2004). Such a difference is related to the morphological make-up of the languages, which affects subject distribution (Roussou & Tsimpli 2006). It may thus have further implications in the interpretation of subjects, allowing more flexible readings in the case of Spanish compared to Italian. In this respect, Roussou and Tsimpli (2006) show that Greek is more permissive with exhibiting VSO orders than Italian. This indicates that Greek resembles more Spanish than Italian regarding VSO, although the latter languages lack case morphology in their D-system, as opposed to Greek. This is a useful observation which helps the comparison between the present study and the results in Sorace et al. (2009).

Returning to the crucial finding in Sorace et al. (2009), the authors also claim that 'learning the efficient coordination of the multiple factors involved in the choice of pronominal forms, and especially sensitivity to redundancy, is a demanding task that requires many years of exposure to be completely acquired and may be particularly taxing for bilingual speakers' (ibid: 474). The finding therefore seems to be also related to the developmental stage of acquisition of their young participants. There is evidence that monolingual and bilingual children, speakers of NS languages, such as Greek, Spanish and Italian, manifest inconsistent pronominal resolution preferences, since sensitivity to the relevant discourse and pragmatic conditions emerges late in the acquisition schedule (see Argyri & Sorace 2007; Sorace et al. 2009; Pinto 2012; Shin & Cairns 2012; Tsimpli 2014; Montrul & Sánchez-Walker 2015; Papadopoulou et

al. 2015; Sorace 2016). Thus, the particular language combination, involving variation at microparametric levels, together with the developmental phase in which the children are in the language acquisition process may explain Sorace et al.'s (2009) crucial finding. Questions arising include whether the bilingual children exhibit the same performance in their other language, i.e. Spanish, as well as whether this linguistic behaviour persists over time and for how long.

Microparametric variation was also found in the present study with respect to the scope of OSP. Contrary to Sorace et al. (2009), the Greek OSP, being a strong pronominal, did not permit language contact effects from Spanish, i.e. another consistent NS language. In other words, the OSP did not manifest the status of the default in the performance of different types of adult bilinguals in Greek. Apart from differences in the distribution of OSP, microparametric variation exists at other levels. As Roussou and Tsimpli (2006) show, it has been attested in VSO orders, which are allowed in both Greek and Spanish, but such a distribution is wider in Greek than in Spanish. In Greek, by virtue of the morphological properties in the grammar, VSO seems to be more productive since it does not depend on topicalisation or focusing, whereas in Spanish it does depend on such functions, as previously mentioned. Comparing the three languages reveals that 'Italian differs from Greek and Spanish in that it does not (readily) allow for VSO' (Roussou & Tsimpli 2006: 321), while Spanish and Greek, although also manifesting differences, seem to be typologically closer in the distribution of VSO than Spanish and Italian. This occurs despite of the fact that both Spanish and Italian have no morphological case on the D-system of their grammars. Consequently, as Roussou and Tsimpli (2006: 332) explain, 'Spanish would be expected to behave like Italian in blocking VSO. However, this is not the case. In fact, Spanish is a cross between Greek (VSO is possible) and Italian (no case morphology on D)'. In sum, the typological similarity between Greek and Spanish in the VSO distribution offers further evidence which suggests that, in bilingualism, influence from the grammatical system of Spanish to Greek is weak compared to influence from Spanish to Italian. The Spanish-Italian pair, apart from microparametric variation in the context of pronouns, also manifests further variation in subject distribution as it instantiates in VSO orders. This difference, therefore, may allow the

implication of weakening or underspecification of formal features more easily than in the Greek-Spanish pair in cases of bilingualism (see Tsimpli et al. 2004; Roussou & Tsimpli 2006; Sorace & Serratrice 2009).

Taken together, the findings of the present and previous research on subjects involving NS languages point to three interrelated facts:

(a) the language combination: a NS language in contact with a non-NS language is expected to trigger (more) language-contact effects than two NS languages in contact (as predicted by the representational account)

(b) microparametric variation: two NS languages are not expected to be identical in subject distribution (Sorace et al. 2009; Roussou & Tsimpli 2006; Filiaci 2011), which may trigger crosslinguistic influence, but this crucially also relates to:

(c) the nature of OSP in a given language, e.g. the Greek third-person OSP was found to differ from the Spanish OSP, although both obey similar pragmatic constraints; the particular nature of the OSP ultimately influences language-contact effects.

The present findings also show the vulnerability of NS, which seem to be influenced by interface conditions in bilingual situations involving two NS languages and to be related to cognitive constraints due to age and/or bilingualism. Overuse of NS involving ambiguity and infelicitous interpretation of NS indicate that this behaviour is generally due to computational demands affecting the efficiency of processing resources in production and interpretation in integrating different levels of linguistic knowledge. Overextension of NS in TS contexts is a pattern which has been attested in previous research with bilinguals, especially in Spanish (e.g. Montrul 2004a; Montrul & Rodríguez Louro 2006; Sorace et al. 2009; Montrul & Sánchez-Walker 2015; Montrul 2016a). However, as Sorace et al. (2009: 474) observe, 'this problem may have been overlooked in the analysis of production data in previous studies since ambiguity is less salient than redundancy'. The present study has provided evidence suggesting that ambiguity may be more salient or problematic than redundancy in different bilingual domains. It also lends support to the claim that crosslinguistic influence does not necessarily result in language attrition, in line with Tsimpli (2017a), which is further discussed in Chapter 9.

8.5 Outline of major findings and conclusions

Summarising the major findings of the present research, Tables 8.1 and 8.2 show the most relevant differences and similarities between the two monolingual groups as revealed in their performance in both production and interpretation. Subsequently, Tables 8.3 and 8.4 display the major findings with regard to the linguistic behaviour of bilinguals. The tables include the predictions and information on whether they were confirmed or not. They also include a concise presentation of the particular findings, a brief outline of possible explanation(s) expounded in the previous sections and the conclusions drawn from the findings.

Table 8.1. Greek vs Spanish monolinguals on OSP

Differences between Greek and Spanish on OSP

Prediction	Greek and Spanish differ in the scope of the OSP.
Confirmed	Yes
Finding	Spanish OSP has wider scope than Greek OSP in production and interpretation.
Possible explanation	The Greek OSP is a strong pronominal with deictic features compared to the Spanish OSP, which is relatively weak, hence the different discourse properties of OSP shown in the two languages.
Conclusion	Crosslinguistic difference in the properties of the two pronominal systems (microparametric variation), thus the scope of OSP varies across the languages, being wider in Spanish and narrower in Greek.

Table 8.2. Greek vs Spanish monolinguals on NS

	Similarities between Greek and Spanish on NS
Prediction	Greek and Spanish resemble in the scope of NS and allow NS in TS.
Confirmed	Yes
Finding	In both languages, NS were felicitously used in TS in production and were interpreted half the time as encoding TS in interpretation.
Possible explanation	NS are used due to economy principles. The felicitous use of NS in TS depends on presence of cues in the (extra)linguistic context.
Conclusion	NS are used in TS, thus involve external interface conditions.

Table 8.3. Bilinguals vs monolinguals on OSP

	Differences between bilinguals and monolinguals on OSP
Prediction	Bilinguals overuse OSP and over-interpret OSP as involving TC.
Confirmed	No
Findings	・ No overgeneralisation of the scope of OSP neither in TC nor in TS in production and interpretation ・ Overuse of LS in TC in production by L2ers and HS
Possible explanations	・ The Greek OSP is resistant to language contact effects and/or crosslinguistic influence due to its discourse properties and the language combination involving another NS language ・ Overuse of LS due to concern for successful communication ・ Potential use of LS as the default (instead of the OSP)
Conclusion	The Greek OSP is not vulnerable at the syntax-discourse/pragmatics interface in contact with Spanish.

Table 8.4. Bilinguals vs monolinguals on NS

	Differences between bilinguals and monolinguals on NS
Prediction	Bilinguals overuse NS in TS and over-interpret NS as involving TS.
Confirmed	Yes
Findings	Production: ・ Older immigrants and HS overuse NS in TS ・ Bilinguals are relatively prone to produce ambiguity Interpretation: ・ Older speakers of Greek, irrespective of linguistic background, link NS to the most recent antecedent (implying TS) ・ HS tend to link NS to the most recent antecedent (implying TS)
Possible explanations	Production: ・ Performance constraints due to bilingualism and to being used to oral discourse, which lead to strategies that favour cognitive economy, resulting in an apparent loss of pragmatic content or restrained performance which entails decreased efficiency in perspective-taking abilities ・ Use of NS as the default Interpretation: ・ General cognitive resources, which favour 'recency of mention' as the preferred ambiguity resolution strategy, enhanced by bilingualism
Conclusion	NS are vulnerable involving external interface conditions.

8.6 Heritage language maintenance: sociolinguistic and linguistic factors

Mobility of people in the context of migration is associated with language change since it entails language contact and consequently the development of societal and individual bi-/multilingualism involving a majority language. Within three generations, this situation usually leads to language shift, which is defined as the piecemeal abandoning of a first/native language in favour of another language (Pauwels 2016), thereby an immigrant language is gradually replaced by a more dominant language. This is also the recurrent scenario regarding the Greek language in Chile. As a result of the gradual language disuse, third generations of original immigrants are normally monolingual or at least 'L2' dominant. In the case of Greek in Chile, such a process is possibly accelerated by the geographical distance between the countries. Language shift mainly concerns the fact that 'if a language ceases to be transmitted domestically, the bedrock of its continuing tradition is undermined' (Coulmas 2013: 161). Therefore, a speech community's temporal continuity does not guarantee neither long-term maintenance nor linguistic homogeneity or uniformity (ibid: 159).

Some macro-social factors involved in a community's ethnolinguistic vitality are related to demography, institutional support and status (see Giles, Bourhis & Taylor 1977). With respect to the broader Greek community in Chile, the absolute size of the group and its relative strength in the total population are quite small. This is intensified by mixed marriages, sporadic continuing migration and provision of few social settings for language use. Moreover, since Greek is not an officially recognised minority language in Chile, the institutional support is confined in the field of education, but predominantly in community-operated (out-of-school) small-scale programmes with rather poor attendance. In regard to status, the group's position in the social prestige hierarchy considerably varies from a socioeconomic point of view. On the other hand, the Chilean society is positively disposed towards the Greek people, their language and culture while Greeks of all generations living in Chile feel proud about their cultural and linguistic heritage, as already mentioned in §1.2. Speaking Greek in Chile thus seems to be a case of 'prestige bilingualism' (Coulmas 2018), yet without involving particular career advantages as in other analogous cases.

It is pertinent to consider the above facts along with the results of the present study, which showed no signs of attrition in the scope of OSP due to language contact. Overextension of the OSP to pragmatically less appropriate contexts is a common observation in several situations of bilingualism (L1 attrition, L2 learning, heritage language acquisition) as shown in related research. The importance of such a finding lies in the fact that it was unpredicted by prevalent theoretical accounts driven by the IH, a hypothesis that has been highly influential in linguistic and psycholinguistic approaches to bilingualism. Potential attrition signs emerged as some simplification of grammatical structures related to use of NS due to processing limitations. Greek seems to be a case of a well-maintained heritage language in the Chilean context, at least in the morphosyntactic domain of OSP, which has been repeatedly shown to be vulnerable across language-contact situations in other studies.

Since Spanish has been 'invasive' in this situation, the sociolinguistic and linguistic conditions for robustness to obtain in a heritage language such as Greek in a remote world zone such as Chile need to be explored. From a sociolinguistic perspective, this point is especially relevant for HS since they largely determine the locus of language change and eventual shift, although it often starts in the parental generation. The fact that competency in Greek does not usually improve employability in Chile, as opposed to e.g. Spanish in the USA, should be also considered since this is a factor that increases motivation for heritage language maintenance.

In contrast to Greek in Chile, Spanish in the USA is not as highly regarded as a heritage language, despite its numerical strength and its historical and geographical importance. The visible manifestation of Spanish in the US social and linguistic landscape and its ensuing utility are at odds with a society which has been characterised as 'anti-bilingual' because of putative threats to the country's national unity (Cashman 2006; Austin et al. 2015). Nonetheless, there is a continuous influx of a large number of new Hispanic immigrants in the USA, presence of Spanish in social institutions, such as mass media and religion, as well as accessibility of some educational support in certain states. These factors contribute to promoting maintenance of Spanish as a minority language at the societal level, although maintenance is not guaranteed at the individual level.

Research in heritage Spanish in the USA provides evidence of impact from English in several morphosyntactic structures since early stages of the acquisition development. In particular, OSP distribution has been shown to manifest significant differences from monolingual norms both quantitatively and discourse-pragmatically (see e.g. Silva-Corvalán 1991, 1993, 2014; Montrul 2004a, 2016a; Montrul & Sánchez-Walker 2015; Montrul 2018; Shin 2018). Thus, language change seems to be more evident in heritage Spanish in the USA than in heritage Greek in Chile, despite the availability of more opportunities for maintenance in the former case than in the latter. Attitudes towards the heritage language and its status as a minority language in the society play a crucial role. From a linguistic perspective, however, a comparable situation obtains in heritage Greek in contact with non-NS languages. Kaltsa et al. (2015) showed that Greek in contact with Swedish gives rise to crosslinguistic effects in OSP interpretation. In L1 attrition, the Greek OSP was affected in contact with Swedish (Kaltsa et al. 2015) and English (Tsimpli et al. 2004) in first-generation immigrants.

The aforementioned studies involving heritage Greek and heritage Spanish in contact with English in relation to the Greek-Spanish pair of this study point to the following linguistic and sociolinguistic factors to be considered in contexts of heritage language maintenance: (a) the typological distance between the languages of the given pair; (b) attitudes towards the language, which are largely determined by 'the value placed on its speakers by the hegemonic group' (Cashman 2006: 42); and (c) the role of formal education and development of literacy skills.

The success of language maintenance of minority/heritage Greek in Chile seems to be related to the typological similarity between Greek and Spanish. In addition, although spoken on a small scale, heritage Greek seems to be preserved thanks to 'the micro-social arena of the family as the agent of spontaneous intergeneration language transmission' (Coulmas 2013: 161). In particular, first-generation immigrants actively and extensively use Greek on a daily basis and are very eager to transmit it to the next generation. This situation is consolidated through positive attitudes and behaviours towards the language and its users by the wider society as well as the Greek speakers themselves, which results in enhancing some kind of language loyalty on the part of the latter. This is also linked to processes of identity negotiation with

the heritage language being the vehicle of cultural inheritance viewed as 'a tangible legacy of the past that looks forward to the future' (Trifonas & Aravossitas 2014: xiii).

Without deviating from the most common scenario, however, third generations rarely or scarcely speak the heritage language. In this study, there was only one third-generation HS with fluent proficiency in heritage Greek. Observably, thus, there is a process of shifting away from the heritage language, which is ultimately abandoned across spheres of usage by third-generation individuals, some of whom deliberately choose to 'reconnect' to heritage/L2 Greek in adulthood. Language shift implies the process and outcome of disappearance of the heritage language 'from the specific speech community (or part thereof) that finds itself in the contact situation' (Pauwels 2016: 18). A relatively swift language shift is certainly what occurs in the case of Greek in Chile, with some family enclaves of resistance (also largely assimilated into the host society) possibly reinforced to some degree by newcomers from the source language community. Language maintenance entails 'situations that exhibit some continuation of L1, however minimal, over an extended period of time' (Pauwels 2016: 21). To the extent that there is some degree of retention of L1/heritage Greek during a period of continued use since the initial language contact, Greek in Chile involves a case of language maintenance, which eventually results in language shift. Nonetheless, for the period of maintenance, Greek seems to be a well-maintained heritage language in this context, arguably due to the typological similarity between the languages in contact and to positive attitudes of the society towards Greek.

This opens up a reflection on educational issues with regard to heritage language maintenance. It is beyond doubt that education helps the development of literacy skills, which in turn, give access to additional opportunities for regular contact with the language. Literacy provides a greater stability in the HS linguistic system, thereby decelerating the process of language attrition, change and shift. Therefore, literacy offers huge potential to foster heritage language maintenance. The multidimensional field of heritage language education is increasingly attracting the attention of several interrelated research fields (e.g. linguistics, pedagogy, policymaking) and relevant issues should be definitely considered in future work.

9. Conclusions

9.1 Summary of the study and conclusions

The present study aimed to investigate the interface phenomenon of third-person subject distribution in NS languages, which has been shown to be problematic in language contact situations. The focus was on two NS languages, namely Greek and Chilean Spanish, in adult speakers in monolingualism and bilingualism. The particular language combination of Greek and Spanish in the context of bilingualism is understudied with regard to the distribution of subjects. The bilingual domains were related to the broad context of migration, with Greek as a heritage/minority language in Chile, and considered first-, second- and third-generation immigrants as well as Chilean L2 speakers of Greek with strong links with the language. The data were obtained from oral production of narratives (Study 1) and interpretation of ambiguous anaphora (Study 2). All elicited data were submitted to statistical analyses while the production data were as well qualitatively scrutinised. The effect of age was examined through logistic regression analyses in all contexts of interest.

The first question of each study involved a direct comparison of Greek and Spanish. The aim was to discover differences between the two languages, which were expected to emerge in the scope of the OSP. The two languages were largely similar, sharing analogous clause structures and displaying generally similar properties on the distribution of subject forms (NS, LS, OSP). The findings, confirming the predictions, showed crosslinguistic differences in the scope of OSP in TS between the two languages due to deictic distinctions, with Greek OSP carrying deictic properties which are less pronounced in its Spanish counterpart. This evidences the fact that NS languages may not be identical with regards to subject distribution. Another relevant key aspect which emerged in examining oral narratives in Greek and Spanish was the felicitous use of NS in TS contexts. NS were also found to be flexible or ambiguous between TC and TS in anaphora resolution in both languages. Thus, NS displayed a more variable distribution than sometimes assumed.

The second question of each study examined Greek in contact with Spanish in Chile focusing on immigrants, HS and L2ers. The aim was to discover differences from monolingual Greek in subject distribution and to disentangle their sources. The bilingual performance in Greek was expected to differ from the monolingual due to language contact, which could enable crosslinguistic influence and/or different processing of subject properties residing at the syntax-discourse/pragmatics interface. Crucially, overextension of the scope of the OSP was anticipated in line with predictions stemming from the IH (Sorace 2011). However, the Greek OSP in contact with Spanish manifested impervious discourse behaviour with no signs of language-contact effects or crosslinguistic influence. This may relate to the properties of the Greek OSP, which is identical in form with the demonstrative and has a special status.

Conversely, overextension of the scope of NS was found in bilinguals and was explained as being due to performance limitations induced by bilingualism, which give rise to cognitively economical processing operations. Specifically, in the domain of L1 attrition, signs of erosion were detected only in advanced stages of attrition, i.e. in first-generation immigrants with lengthier periods of time in the host country, who were moreover of advanced age. Relatively frequent use of NS and instances of ambiguity were observed in older immigrants, who also employed the 'recency' strategy in anaphora resolution. This preference was similar in older Greek monolinguals, suggesting that general cognitive limitations affect older speakers, irrespective of linguistic background, in their Greek performance and this is possibly intensified by bilingualism. As regards HS, their performance was comparable to that of older immigrants in the misuse of NS. This finding points to a tendency towards minimisation of linguistic forms in speakers who are more sensitive to cognitive load due to bilingualism and/or age-related cognitive decline, as well as potential input effects. Ambiguity, emerging in all bilingual groups, including L2 speakers, indicates that the area of NS in TS is vulnerable to instability. In conclusion, the findings of this study support assertions that language attrition is generally minimal and apparently only temporary rather than permanent to the extent that the overextension of NS to TS contexts seems to be triggered by processing constraints.

The findings confirm the representational account (Tsimpli et al. 2004), which predicts no attrition in cases of combinations of languages that instantiate a similar pragmatic setting. In line with Tsimpli (2017a), this study also lends support to the claim that language contact and bilingualism do not necessarily entail language attrition (see also Flores 2017). Furthermore, 'attrition' can be also found in monolingual speakers, as shown in Kaltsa et al. (2015) and corroborated in the present findings, thereby involving cases of age-related language attrition, which is clearly not due to language contact. In both studies, older monolinguals followed economy strategies more often than younger speakers. Thus, as argued by Tsimpli (2017a: 761), crosslinguistic influence and L1 attrition involve two conceptually and empirically distinct effects: 'the former is the inevitable result of bilingualism, while the latter is neither exclusively nor necessarily so'.

The findings of the present study also lend support to the VH (Prada Pérez 2018), which predicts that more categorical distributions (i.e. the OSP in Greek) are less vulnerable than more variable distributions (i.e. NS in Greek). Lastly, L2ers and HS also tended to overuse LS in TC, suggesting that such over-expliciteness may indicate use of LS as the default form instead of the OSP in production of narratives in the weak language.

9.2 Implications on broader issues

In light of the findings, it can be concluded that pronominal use in Greek in bilingual speakers is not exactly monolingual-like but it is not fundamentally different either. As Sorace (2016: 11) observes, 'bilinguals tend to make more extensive use of an option that monolinguals also employ'. In this study, however, the overextended option was the NS and not the OSP implied by Sorace. Only the scope of NS was found to be variable, bringing forward its linguistic complexity, which can be arguably viewed as arising from external interface conditions and possibly involving a default use. According to Sorace (2016: 11), 'the differences between monolinguals and bilinguals in language processing, like the differences in general cognition, can be seen as advantageous or disadvantageous only if one takes the monolingual system as a point of reference'. However, if such differences involve ambiguity caused by

insufficiently informative constructions, which may compromise the quality of communication, such behaviour may be regarded as disadvantageous beyond comparisons with monolingual norms. This stands in stark contrast to overuse of overt forms, which do not affect recoverability in the discourse.

Pragmatically appropriate use and interpretation of referential subjects in discourse is a key aspect for efficient communication. 'By definition, coherent discourse, with its multiple recurring strands, makes language-coded information mentally more accessible, memorable and retrievable' (Givón 2016: 30). Academic work may have valuable applications in the field of monolingual, bilingual and L2 education. For instance, research findings can indicate linguistic areas shown to be persistently vulnerable in bilingual situations as well as the identified differences between L2/bilingual speakers and monolinguals in these areas. 'For pedagogical purposes, explaining how and why differences obtain is of little use [...], but knowing what those differences are can be very useful' (Rothman, Tsimpli & Pascual y Cabo 2016: 22). In any case, from an educational point of view, in line with Montrul (2016b), the focus should be on bilingualism, its advantages and its maintenance, instead of sticking to differences with monolingual norms (see also Polinsky & Kagan 2007; Grosjean 2008; Cummins 2014; Trifonas & Aravossitas 2014; Tsimpli 2017b).

9.3 Limitations and future work

It is hoped that the present study contributed to current debates revolving around differences between NS languages and the bilingual behaviour respecting the use of subjects in production and interpretation. It is certainly necessary to conduct further research in order to corroborate or refute the conclusions reached. In addition, there are methodological facts which should be recognised as limitations of the study.

Due to the nature of the methodological approach, the research participants were not tested in exactly the same circumstances, i.e. the specific environment was not controlled. Although this could be seen as a potential drawback, it is regarded as an advantage since the naturalistic settings chosen by the participants contributed to ecologically approaching the conditions in which the speakers use the language.

Further, a working-memory task should have been included to support the potential age-related weakening explanation, but for practical reasons this was unfeasible.

In regard to production, the picture-based narrative task was presented to the participants with the instruction to perform it for an imaginary listener who could not see the pictures. Although it is deemed that the participants, as adults, complied with this agreement, the possibility that some of them assumed a degree of shared knowledge (especially the older ones) cannot be ruled out. The annotation of narratives was also carried through by one evaluator, namely myself. In related research, ideally two readers scrutinise the data independently in order to assure reliability of the analysis (e.g. Montrul & Rodríguez-Louro 2006; Lubbers Quesada & Blackwell 2009; Hendriks et al. 2014; Schmitz & Scherger 2017). Although I worked thoroughly on the production data for more than two years, often reconsidering the original recordings along with transcriptions, it should be recognised that two sets of eyes would have increased reliability. This, however, was not possible in the context of the present study since, as a PhD research, it was expected to be fulfilled exclusively by myself in all phases.

A wealth of linguistic and sociolinguistic information was collected from the biographical interviews of the bilingual participants. The linguistic data could serve to corroborate findings obtained from the narratives, considering also the level of individual performance. This task is left for future work since the focus on narratives entailed a more controlled use of vocabulary, thus allowing more effective group comparisons at least in the first instance. Sociolinguistic aspects of the participants' profiles, such as past experiences with the input, ways in which they maintain bilingualism and individual differences, could be also exploited further in relation to their linguistic performance. Due to space constraints, only the primary sociolinguistic information of the participants was included in the relevant chapter and appendices.

While the focus was on subjects, the elicited narratives lend themselves particularly well to investigating reference in the discourse more generally, i.e. considering also maintenance or reintroduction of referents with other syntactic functions (e.g. objects and prepositional phrases). Furthermore, subject position with respect to verb is a

topic to be examined in future work using the same monolingual and bilingual data. Other aspects to be considered at later stages include the focused sentences and the subject-headed relative clauses, which were excluded from the present analyses.

Although it would have been optimal to also examine the bilinguals' performance in Spanish, not only in Greek, so to allow a fuller perspective through a holistic view of bilingualism (Grosjean 2008; Montrul 2016b), this was not possible due to practical reasons. The bilingual production data collected in Spanish were not included in the study because, in my view, lacked optimal quality since they were elicited on the same day when data collection in Greek took place. The speakers were shown to be biased towards their Greek version of the stories. It was then difficult in practical terms to rearrange meetings with all or most of the participants. Thus, it was decided that the particular approach to bilinguals would be unilingual. The question, however, of whether interface processing problems caused by bilingualism are mirrored in both languages has been left open and should be definitely addressed.

A related question to be explored in future work is the comparison of the bilinguals of the present study with their mirror image of Spanish-Greek bilinguals in Greece, namely first generation of Chilean immigrants, HS and L2 speakers with family or personal connections with Chilean Spanish. This would offer important insights into crosslinguistic interaction, bilingualism effects and the role of the majority language in the same language combination in the context of migration but in the mirroring host country.

References

Abbott, B. (2006). Definite and Indefinite. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics,* 2nd Edition, Vol. 3 (pp. 392-399). Oxford: Elsevier.

Abrams, L. & Farrell, M. (2010). Language processing in normal aging. In J. Guendouzi, F. Loncke & M. J. Williams (Eds.), *The Handbook of Psycholinguistic and Cognitive Processes. Perspectives in Communication Disorders* (pp. 49-73). New York: Routledge.

Ackema, P. & Neeleman, A. (2017a). Default person versus default number in agreement. In M. Reeve, L. Franco & M. Moreno (Eds.). *Non-local dependencies in the nominal and verbal domain*. (Open Generative Syntax). Language Science Press.

Ackema, P. & Neeleman, A. (2017b). *Features of Person*. Linguistic Inquiry Monographs. MIT Press.

Adger, D. (2003). *Core Syntax. A minimalist approach*. Oxford & New York: Oxford University Press.

Alexiadou, A. & Anagnostopoulou, E. (1998). Parametrizing AGR: word order, V-movement and EPP checking. *Natural Language and Linguistic Theory*, 16, 491-539.

Alexiadou, A., Haegeman, L. & Stavrou, M. (2007). *Noun phrase in the generative perspective*. Berlin, New York: Mouton de Gruyter.

Altman, G. T. M. & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, 57, 502-518.

Alonso-Ovalle, L., Clifton, C., Frazier, L. & Fernandez-Solera, S. (2002). Null vs. overt pronouns and the topic–focus articulation in Spanish. *Journal of Italian Linguistics*, 14, 151-169.

Almor, A. (1999). Noun-Phrase Anaphora and Focus: The Informational Load Hypothesis. *Psychological Review*, 106(4), 748-765.

Almor, A. (2000). Constraints and Mechanisms in Theories of Anaphor Processing. In M. W. Crocker, Pickering, M. & Clifton, C. (Eds.), *Architectures and mechanisms for language processing* (pp. 341-353). New York: Cambridge University Press.

Almor, A. & Nair, V. A. (2007). The Form of Referential Expressions in Discourse. Language and Linguistics Compass, 1(1-2), 84-99.

Andreou, M. (2015). *The effects of bilingualism on verbal and non verbal cognition: The micro- and macro-structures of narratives in the weak and the dominant language of the bilingual child*. PhD dissertation, Aristotle University of Thessaloniki.

Arbuckle, T. Y., Nohara-LeClair, M. & Pushkar, D. (2000). Effect of off-target verbosity on communication efficiency in a referential communication task. *Psychology & Aging*, 15, 65-77.

Ariel, M. (1990). *Accessing noun-phrases antecedents*. London & New York: Routledge

Ariel, M. (1994). Interpreting anaphoric expressions: A cognitive versus a pragmatic approach. *Journal of Linguistics,* 30, 3-42.

Ariel, M. (2001). Accessibility theory: An overview. In T. Sanders, J. Schliperoord & W. Spooren (Eds.), *Text representation* (pp. 29-87). Human cognitive processing series. John Benjamins.

Ariel, M. (2012). Research paradigms in pragmatics. In K. Allan & K. Jaszczolt (Eds.), *The Cambridge Handbook of Pragmatics* (pp. 12-45). Cambridge, UK: Cambridge University Press.

Ariel, M. (2017). Pragmatics and Grammar: More pragmatics or more grammar. In Y. Huang (Ed.), *The Oxford Handbook of Pragmatics* (pp. 471-792). Oxford: Oxford University Press.

Argyri, E. & Sorace, A. (2007). Crosslinguistic influence and language dominance in older bilingual children. *Bilingualism: Language and Cognition*, 10(1), 77-99.

Arnold, J. E., Bennetto, L., & Diehl, J. J. (2009). Reference production in young speakers with and without autism: Effects of discourse status and processing constraints. *Cognition*, 110(2), 131-146.

Arnold, J. E. & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, 56(4), 521-536.

Arnold, J. E., Kaiser, E., Kahn, J. M. & Kim, L. K. (2013). Information structure: linguistic, cognitive, and processing approaches. *WIREs Cognitive Science*, 4, 403-413.

Ayoun, D. (2003). *Paremeter Setting in Language Acquistion*. London & New York: Continuum.

Austin, J., Blume, M. & Sánchez, L. (2015). *Bilingualism in the Spanish-Speaking World. Linguistic and Cognitive Perspectives*. Cambridge, UK: Cambridge University Press.

Baker, C. (2011). *Foundations of Bilingual Education and Bilingualism*. 5th Edition. Clevedon UK: Multilingual Matters.

Bel, A. (2003). The syntax of subjects in the acquisition of Spanish and Catalan. *Probus* 15, 1-26.

Bel, A. & García-Alcaraz, E. (2015). Subject pronouns in the L2 Spanish of Moroccan Arabic speakers. In T. Judy & S. Perpiñán (Eds.), *The Acquisition of Spanish in Understudied Language Pairings* (pp. 201-232). Amsterdam: John Benjamins.

Belletti, A. (2004). Aspects of the low IP area. In L. Rizzi (Ed.), *The Structure of IP and CP. The Cartography of Syntactic Structures* (pp. 16-51). Oxford: Oxford University Press.

Belletti, A., Bennati, E. & Sorace, A. (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. *Natural Language & Linguistic Theory*, 25, 657-689.

Benmamoun E., Montrul S. & Polinsky M. (2010). *Prolegomena to Heritage Linguistics*. White Paper. Harvard University.

Benmamoun, E. Montrul, S. & Polinsky, M. (2013). Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics*, 39, 129-181.

Beys, M. (2009). Subject-to-Object Raising Constructions in Modern Greek. In *Selected Papers from the 18th International Symposium on Theoretical and Applied Linguistics* (pp. 105-113). Thessaloniki: Monochromia Publishing.

Biberauer, T. Holmberg, A. Roberts, I. & Sheehan, M. (2010). *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press.

Bini, M. (1993). La adquisición del italiano: Más allás de las propiedades sintácticas del parámetro pro-drop. In J. M. Liceras (Ed.), *La lingüística y el análisis de los sistemas no nativos* (pp. 126-139). Ottawa: Dovehouse.

Blackwell, S. E. (2003). *Implicatures in discourse: The case of Spanish NP anaphora*. Amsterdam: John Benjamins.

Blackwell, S. E. & Quesada, M. L. (2012). Third-person subject reference in native speakers' and L2 learners' narratives: Testing (and revising) predictions of the Givenness Hierarchy for Spanish. In K. Geeslin & M. Díaz-Campos (Eds.), *Proceedings from the Hispanic Linguistics Symposium 2010*, Somerville, MA: Cascadilla Proceedings Project.

Borer, H. (1986). I-Subjects. *Linguistic Inquiry*, 17, 375-416.

Burgos Cacharos, S. (2009). *Colectividad Helénica de Santiago, 1918-2006*. Santiago: Lom Ediciones.

Callahan, S. M., Nicol, J., Love, T. & Swinney, D. (2007). *All pronouns are not created equal: The processing and interpretation of null and overt pronouns in Spanish*. Poster presented at the 20th Annual CUNY Sentence Processing Conference, La Jolla, Calif.

Camacho, J. A. (2013). *Null Subjects*. Cambridge Studies in Linguistics 137. Cambridge, UK: Cambridge University Press.

Cardinaletti, A. (2014). Cross-linguistic variation in the syntax of subjects. In M. C. Picallo (Ed.), *Linguistic Variation in the Minimalist Framework*. Oxford Scholarship online. DOI: 10.1093/acprof:oso/9780198702894.001.0001

Cardinaletti, A. & Starke, M. (1999). The Typology of Structural Deficiency. A Case Study of the Three Classes of Pronouns. In H. van Riemsdijk (Ed.), *Clitics in the Languages of Europe* (pp. 145-233), Mouton-De Gruyter, Berlin.

Carminati, M. N. (2002). *The processing of Italian subject pronouns*. PhD dissertation, University of Massachusetts Amherst.

Carminati, M. N. (2005). Processing reflexes of the Feature Hierarchy (Person > Number > Gender) and implications for linguistic theory. *Lingua*, 115, 259-285.

Cashman, H. R. (2006). Who wins in research on bilingualism in an anti-bilingual state? *Journal of Multilingual and Multicultural Development*, 27(1), 42-60.

Centre for the Greek Language. (2012). *Guide to the Examination for the Certificate of Attainment in Greek*. Ministry of Education, Lifelong learning and Religious affairs, Centre for the Greek Language.

Chamorro, G., Sorace, A. & Sturt, P. (2015). What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition. *Bilingualism: Language and Cognition*,19(3), 520-532.

Chamorro, G. & Sorace, A. (2018). The Interface Hypothesis as a framework for studying L1 attrition. To appear in B. Köpke, & M. Schmid (Eds.), *The Oxford Handbook of Language Attrition*. Oxford: Oxford University Press.

Chatzidaki, A. (2016). Η ελληνική γλώσσα ως δεύτερη και ως ξένη στη διασπορά. Εκπαιδευτικές δράσεις και νέα δεδομένα. In E. Skourtou & V. Kourti-Kazoulli (Eds.), *Διγλωσσία και Διδασκαλία της Ελληνικής ως Δεύτερης Γλώσσας* (pp. 70-95). Athens: Kallipos.

Chatzikyriakidis, S., Michelioudakis, D. & Spathas, G. (2015). Greek focus operators and their associates. In *Studies in Greek Linguistics 35, Proceedings Annual Meeting of the Department of Linguistics* (pp. 167-179), School of Philology, Aristotle University of Thessaloniki 8-10 May 2014. Institute of Modern Greek Studies, Conference paper.

Chiou, M. (2012). The meaning of going marked: neo-Gricean pragmatic analysis of the alternate set {zero vs. full pronoun/anaphor} in Modern Greek. *Γλωσσολογία/Glossologia*, 20, 39-53.

Chiriacescu, S. I. (2014). *The discourse structuring potential of indefinite noun phrases. Special markers in English, German and Romanian*. PhD dissertation, University of Stuttgart.

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, 36, 1-22.

Clements, M. & Domínguez, L. (2016). Reexamining the acquisition of null subject pronouns in a second language. *Linguistic approaches to bilingualism*, 1-25.

Coulmas, F. (2018). *An Introduction to Multilingualism: Language in a Changing World*. Oxford: Oxford University Press.

Coulmas, F. (2013). *Sociolinguistics: The Study Of Speakers' Choices*. 2nd Edition. Cambridge, UK: Cambridge University Press.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Press Syndicate of the University of Cambridge.

Council of Europe. (2018). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989

Cummins, J. (2014). Mainstreaming Plurilingualism: Restructuring Heritage Language Provisions in Schools. In P. P. Trifonas & T. Aravossitas (Edis.), *Rethinking heritage language education* (pp. 1-19). Cambridge, UK: Cambridge University Press.

Damanakis, M. (2007). *Ταυτότητες και εκπαίδευση στη διασπορά*. Gutengerg.

Dimitriadis, A. (1996). When pro-drop languages don't: Overt pronominal subjects and pragmatic inference. In *Proceedings of the 32nd Meeting of the Chicago Linguistics Society. Chicago*, 33-47.

Domínguez, L. (2013). *Understanding Interfaces. Second language acquisition and first language attrition of Spanish subject realization and word order variation*. Amsterdam & Philadelphia: John Benjamins.

Domínguez, L. & Arche, M. J. (2014). Subject inversion in non-native Spanish. *Lingua*, 145, 243-265.

Duarte, M. E. & Soares da Silva, H. (2016). Microparametric variation in Spanish and Portuguese. The null subject parameter and the role of the verb inflectional paradigm. In M. A. Kato & F. Ordóñez, *The Morphosyntax of Portuguese and Spanish in Latin America*. Oxford University Press.
DOI: 10.1093/acprof:oso/9780190465889.001.0001

Ellis, N. C. (1998). Emergentism, Connectionism and Language Learning. *Language Learning*, 48(4), 631-664.

Enríquez, E. V. (1986). La presencia de los pronombres personales sujetos en el mundo hispánico. Estudio comparativo. *Anuario de Letras*, 24(3), 47-70.

Filiaci, F. (2010). Null and overt subject biases in Spanish and Italian: a cross-linguistic comparison. In C. Borgonovo, M. Español-Echevarría, P. Prévost (Eds.), *Selected Proceedings of the 12th Hispanic Linguistic Symposium* (pp. 171-182). Somerville, MA: Cascadilla Proceedings Project.

Filiaci, F. (2011). *Anaphoric preferences of null and overt subjects in Italian and Spanish: A cross-linguistic comparison*. PhD dissertation, The University of Edinburgh.

Filiaci, F., Sorace, A. & Carreiras, M. (2013): Anaphoric biases of null and overt subjects in Italian and Spanish: a cross-linguistic comparison, *Language and Cognitive Processes*, 1-19.

Flores, C. (2008). *A competência sintáctica de falantes bilingues luso-alemães regressados a Portugal. Um Estudo sobre Erosão Linguística*. PhD dissertation, Universidade do Minho.

Flores, C. (2017). Problematizing the scope of language attrition from the perspective of bilingual returnees. *Linguistic Approaches to Bilingualism*, 7(6), 691-695.

Frana, I. (2017). The Role of Discourse Prominence in the Resolution of Referential Ambiguities Evidence from co-reference in Italian. *International Journal of Linguistics*, 9(3), 205-225.

Frascarelli, M. (2007) Subjects, topics and the interpretation of referential *pro*: An interface approach to the linking of (null) pronouns. *Natural Language and Linguistic Theory*, 25(4), 691-734.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Valimaa, T., Baleiuniene, I. et al. (2015). Assessment of narrative abilities in bilingual children. In S. Armon-Lotem, J. de Jong & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 241-274). Bristol: Multilingual Matters.

Gagarina, N., Klop, D., Tsimpli, I. M. & Walters, J. (2016). Introduction. Narrative abilities in bilingual children. *Applied Psycholinguistics*, 37, 11-17.

Gelormini-Lezama, C. & Almor. A. (2011). Repeated names, overt pronouns, and null pronouns in Spanish. *Language and cognitive processes*, 26(3), 437-454.

Gelormini-Lezama, C., & Almor, A. (2014). Singular and Plural Pronominal Reference in Spanish. *Journal of Psycholinguistic Research*, 43(3), 299-313.

Genesee, F. (2016). Shifting Perspectives on Bilingualism. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism Across the Lifespan* (pp. 9-19). De Gruyter Mouton.

Georgopoulos, A. (2017). Anaphora resolution in the interlanguage of Greek and English learners of Spanish: A corpus study. *Studies in Greek Linguistics*, 37, 239-252.

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68, 1-76.

Gijn, R. (2016). Switch reference. An overview. In R. Gijn & J. Hammond (Eds.), *Switch Reference 2.0*. Amsterdam & Philadelphia: John Benjamins.

Giles, H., Bourhis, R. & Taylor, D. M. (1977). Towards a theory of language in ethnic group relations. In H. Giles (Ed.), *Language, Ethnicity and Intergroup Relations*. London: Academic Press, 307-348.

Givón, T. (1983). Topic continuity in discourse: The functional domain of switch-reference. In J. Haiman & P. Munro (Eds.), *Switch-reference and universal grammar*, vol. 2 (pp. 51-82). Amsterdam & Philadelphia: John Benjamins.

Givón, T. (2016). *The Story of Zero*. Amsterdam: John Benjamins Publishing Company.

Goodluck, H., Terzi, A. & Chocano Díaz, G. (2001). The acquisition of control crosslinguistically: structural and lexical factors in learning to licence PRO. *Journal of Child Language*, 28, 153-172.

Grosjean, F. (2008). *Studying Bilinguals*. Oxford: Oxford University Press.

Grosz, B. J., Joshi, A. K. & Weinstein, S. (1995). Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2), 203-225.

Gundel, J. K. (1996). Relevance theory meets the Givenness Hierarchy: An account of inferrables. In T. Fretheim & J. K. Gundel (Eds.), *Reference and Referent Accessibility* (pp. 141-153). Amsterdam & Philadelphia: John Benjamins.

Gundel, J. K., Hedberg, N. & Zacharski, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, vol. 69, no. 2, 274-307.

Gürel, A. (2006). L2 acquisition of pragmatic and syntactic constraints in the use of overt and null subject pronouns. *Inquiries in linguistic development*, 259-282.

Haegeman, L. (1994). *Introduction to Government and Binding Theory*. Oxford: Blackwell.

Hendriks, H. (2003). Using nouns for reference maintenance: A seeming contradiction in L2 discourse. In A. G. Ramat (Ed.), *Typology and second language acquisition* (pp. 291-326). Berlin: Walter de Gruyter.

Hendriks, P., Koster, C. & Hoeks, J. (2014). Referential choice across the lifespan: why children and elderly adults produce ambiguous pronouns. *Language, Cognition and Neuroscience*, 29(4), 391-407.

Hickmann, M. (2003). *Children's Discourse: Person, Space and Time Across Languages*. Cambridge, UK: Cambridge University Press.

Holton, D., Mackridge, P. & Philippaki-Warburton, I. (2002). *Greek. A Comprehensive Grammar,* 2nd ed. New York: Routledge.

Huang, Y. (2014). *Pragmatics*. 2nd Edition. New York: Oxford University Press.

Huang, Y. (2017). Neo-Gricean pragmatics. In Y. Huang (Ed.), *The Oxford Handbook of Pragmatics* (pp. 47-78). Oxford: Oxford University Press.

Hyams, N., Mateu, V., Ortfitelli, R., Putnam, M., Rothman, J. & Sánchez, L. (2015). Parameters in language acquisition and language contact. In A. Fábregas, J. Mateu, J. & M. Putnam (Eds.), *Contemporary Handbook of Linguistic Parameters* (pp. 353-373). London: Bloomsbury.

Hyams, N. & Wexler, K. (1993). On the Grammatical Basis of Null Subjects in Child Language. *Linguistic Inquiry*, 24(3), 421-459.

Iatridou, S. (1993). On Nominative Case Assignment and a few related things. *MIT Working Papers in Linguistics*, 19.

Iraola, M. (2015). *Anaphora Resolution in Children and Adults. An Experimental Study of Mature Speakers and Learners of Basque*. Tübingen: Narr Verlag.

Iverson, M. B. (2012). *Advanced language attrition of Spanish in contact with Brazilian Portuguese*. PhD dissertation, University of Iowa.

Jaeggli, O. & Safir, K. (1989). *The null subject parameter*. Dordrecht: Foris.

Jegerski, J., VanPatten, B. & Keating, G. D. (2011). Cross-linguistic variation and the acquisition of pronominal reference in L2 Spanish. *Second Language Research*, 33(2), 193-221.

James, L. E., Burke, D. M., Austin, A. & Hulme, E. (1998). Production and perception of verbosity in young and older adults. *Psychology & Aging*, 13, 355-367.

Judy, T. (2015). Knowledge and processing of subject-related discourse properties in L2 near-native speakers of Spanish, L1 Farsi. In T. Judy & S. Perpiñán (Eds.), *The Acquisition of Spanish in Understudied Language Pairings* (pp. 169-199). Amsterdam: John Benjamins.

Kaltsa, M., Tsimpli, I. M. & Rothman, J. (2015). Exploring the source of differences and similarities in L1 attrition and heritage speaker competence: evidence from pronominal resolution. *Lingua*, 164, 266-288.

Kaiser, E. & Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and cognitive processes*, 23(5), 709-748.

Keating, G. D., VanPatten, B. & Jegerski, J. (2011). Who was walking on the beach? Anaphora resolution in Spanish heritage speakers and adult second language learners. *Studies in Second Language Acquisition*, 33(2), 193-221.

Kibrik, A. (2011). *Reference in Discourse*. New York: Oxford University Press.

Kiss, K. É. (1995). Discourse configurational languages. Introduction. In K.É. Kiss (Ed.), *Discourse Configurational Languages*. Oxford University Press, Oxford.

Kupisch, T. (2013). A new term for a better distinction? A view from the higher end of the proficiency scale. *Theoretical Linguistics*, 39(3-4), 203-214.

Kupisch, T. & Rothman, J. (2016). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 1-19.

Lafond, L., Hayes, R., & Bhatt, R. (2001). Constraint demotion and null subjects in Spanish L2 acquisition. In J. Camps & C. Wiltshire (Eds.), *Romance Syntax, Semantics and L2 Acquisition* (pp. 121-135). Amsterdam: John Benjamins.

Lardiere, D. (2011). Who is the Interface Hypothesis about? *Linguistic Approaches to Bilingualism*,1(1), 48-53.

Leclercq, P. & Lenart, E. (2013). Discourse Cohesion and Accessibility of Referents in Oral Narratives: A Comparison of L1 and L2 Acquisition of French and English. *Discours* [Online], DOI: 10.4000/discours.8801

Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge: MIT Press.

Liceras, J., Alba de la Fuente, A. & Martínez Sanz, C. (2010). The distribution of null subjects in non-native grammars: Syntactic markedness and interface vulnerability. In *Proceedings of the Mind/Context Divide Workshop* (pp. 84-95). Somerville, MA: Cascadilla Proceedings Project.

Liceras, J. & Fernández Fuertes, R. (2017). Subject omission/production in child Bilingual English and child bilingual Spanish: the view from linguistic theory. *Probus*. DOI 10.1515/probus-2016-0012.

Lipski, J. M. (1994). *Latin American Spanish*. London: Longman.

Lipski, J. M. (2012). Geographical and Social Varieties of Spanish: An Overview. In J. I. Hualde, A. Olarrea & E. O'Rourke. *The Handbook of Hispanic Linguistics* (pp. 1-26). Blackwell Publishing Ltd.

Lohndal, T. & Uriagereka, J. (2016). Third-Factor Explanations and Universal Grammar. In I. Roberts (Ed.), *The Oxford Handbook of Universal Grammar* (pp. 1-18). Oxford Handbook Online.

Lozano, C. (2002). Knowledge of expletive and pronominal subjects by learners of Spanish. *International Journal of Applied Linguistics*, 135-136, 37-60.

Lozano, C. (2006a). *Explaining the "syntax-before-discourse" phenomenon: Pronominal subject distribution in L1 Greek-L2 Spanish*. Paper presented at The Romance Turn II, University of Utrecht, 7th September.

Lozano, C. (2006b). Focus and split intransitivity: the acquisition of word order alternations in non-native Spanish. *Second Language Research*, 22, 1-43.

Lozano, C. (2009). Selective deficits at the syntax-discourse interface Evidence from the CEDEL2 corpus. In N. Snape, Y. I. Leung & M. S. Smith (Eds.), *Representational deficits in SLA* (pp. 127-166). Amsterdam & Philadelphia: John Benjamins.

Lozano, C. (2016). Pragmatic principles in anaphora resolution at the syntax-discourse interface. Advanced English learners of Spanish in the CEDEL2 corpus. In M. Alonso Ramos (Ed.), *Spanish Learner Corpus Research: Current Trends and Future Perspectives* (pp. 236-265). Amsterdam: John Benjamins.

Lozano, C. (2018). The development of anaphora resolution at the syntax-discourse interface: pronominal subjects in Greek learners of Spanish. *Journal of Psycholinguistic Research*, 47(2), 411-430.

Lyons, C. (1999). *Definiteness*. Cambridge, UK. Cambridge University Press.

Lubbers Quesada, M. (2015). *The L2 Acquisition of Spanish Subjects. Multiple Perspectives*. Series Studies on Language Acquisition. Boston/Berlin: De Gruyter.

Lubbers Quesada, M. & Blackwell, S. E. (2009). The L2 acquisition of null and overt Spanish subject pronouns: A pragmatic approach. In J. Collentine, M. E. Garcia, B. Lafford & F. Marcos-Marin (Eds.), *Proceedings of the 2007 Hispanic Linguistics Symposium* (pp. 117-130). Somerville, MA: Cascadilla Proceedings Project.

Luján, M. (1999a). A unified approach to control and obviation. In J. Franco, A. Landa & J. Martín (Eds.), *Grammatical Analyses in Basque and Romance Linguistics* (pp. 105-130). Amsterdam & Philadelphia: John Benjamins.

Luján, M. (1999b). Expresión y omisión del pronombre personal. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (pp. 1275–1316). Madrid: Espasa Calpe.

Manolessou, I. (2001). The Evolution of the Demonstrative System in Greek. *Journal of Greek Linguistics*, 2, 119-148.

Margaza, P. & Bel, A. (2006). Null subjects at the syntax-pragmatics interface: Evidence from Spanish interlanguage of Greek speakers. In G. O'Brien, C. Shea & Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)* (pp. 88-97). Somerville MA: Cascadilla Press.

Mavrogiorgos, M. (2010). *Clitics in Greek. A minimalist account of proclisis and enclisis*. Amsterdam & Philadelphia: John Benjamins.

Mastropavlou, M., Katsiperi, M., Fotiadou, G., Fleva, E., Peristeri, E. & Tsimpli, I. M. (2014). *The Role of Definiteness in Anaphora Resolution. Experimental Psycholinguistisc*, Madrid 1-3 October 2014.

Mayol, L. (2012). An account of the variation in the rates of overt subject pronouns in Romance. *Spanish in Context*. 9(3), 420-442.

Milroy, L. (1987). *Language and Social Networks*. New York: Blackwell.

Miltsakaki, E. (2001). Centering in Greek. In *Proceedings of 15th International Symposium on Theoretical and Applied Linguistics*. Aristotle University, Thessaloniki.

Miltsakaki, E. (2007) A rethink of the relationship between salience and anaphora resolution. In A. Branco, T. McEnery, R. Mitkov, F. Silva (Eds.), *Proceedings of the 6th Discourse Anaphora and Anaphor Resolution Colloquium* (DAARC 2007). Centro de Linguística de Universidade do Porto, Porto.

Miltsakaki, E. (2011). Not All Subjects Are Born Equal: A Look at Complex Sentence Structure. In E. Gibson & N. Perlmutter (Eds.), *The Processing and Acquisition of Reference* (pp. 355-380). Cambridge, MA: MIT Press.

Montrul, S. (2004a). Subject and object expression in Spanish heritage speakers: A case of morpho-syntactic convergence. *Bilingualism: Language and Cognition*, 7(2), 125-142.

Montrul, S. (2004b). *The Acquisition of Spanish. Morphosyntactic Development in Monolingual and Bilingual L1 Acquisition and in Adult L2 Acquisition*. Amsterdam: John Benjamins.

Montrul, S. (2005). Second language acquisition and first language loss in adult early bilinguals: exploring some differences and similarities. *Second Language Research*, 21(3), 199-249.

Montrul, S. (2008). *Incomplete acquisition in bilingualism: re-examining the age factor*. Amsterdam & Philadelphia: John Benjamins.

Montrul, S. (2011). Multiple interfaces and incomplete acquisition. *Lingua*, 121, 591-604.

Montrul, S. (2014). Age of acquisition interactions in bilingual lexical access: A study of the weaker language of L2 learners and heritage speakers. *International Journal of Bilingualism*, 18(3), 274-303.

Montrul, S. (2016a). Heritage language development: Connecting the dots. *International Journal of Bilingualism*, 1-17.

Montrul, S. (2016b). *The acquisition of heritage languages*. Cambridge, UK: Cambridge University Press.

Montrul, S. (2017). Developmental continuity in morphosyntactic attrition. *Linguistic Approaches to Bilingualism*, 7(6), 739 -743.

Montrul, S. (2018). Morphology, syntax, and semantics in Spanish as a heritage language. In K. Potowski (Ed.), *The Routledge Handbook of Spanish as a Heritage Language* (pp. 145-163). London & New York: Routledge.

Montrul, S. & Polinsky, M. (2011). Why not heritage speakers? *Linguistic Approaches to Bilingualism*,1(1), 58-62.

Montrul, S. & Rodríguez-Louro, C. (2006). Beyond the syntax of the Null Subject Parameter: a look at the discourse-pragmatic distribution of null and overt subjects by L2 learners of Spanish. In: Torrens, V., Escobar, L. (Eds.), *The Acquisition of Syntax in Romance Languages* (pp. 401-418). Amsterdam: John Benjamins.

Montrul, S. & Sánchez-Walker, N. (2015). Subject expression in school-age Spanish-English bilingual children. In A. Carvalho, R. Orozco & N. Shin (Eds.), *Subject pronoun expression in Spanish: A crossdialectal perspective* (pp. 231–248). Washington DC: Georgetown University Press.

Moreno Fernández, F. (2014). *La lengua española en su geografía. Manual de dialectología hispánica*. 2nd ed. Barcelona: Ariel.

O'Grady, W. (2008). The emergentist program. *Lingua*, 118, 447-464.

Ordóñez, F. (2012). Clitics in Spanish. In J. I. Hualde, A. Olarrea & E. O'Rourke. *The Handbook of Hispanic Linguistics* (pp. 423-451). Blackwell Publishing Ltd.

Ordóñez, F. & Olarrea, A. (2006). Microvariation in Caribbean/non Caribbean Spanish interrogatives. *Probus*, 18(1), 59-96.

Otheguy, R. & Zentella, A. C. (2012). *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. Oxford University Press.

Otheguy, R., Zentella, A. C. & Livert, D. (2007). Language and dialect contact in Spanish in New York: Towards the formation of a speech community. *Language*, 83, 770-802.

Oxford Reference (2017). 3rd Edition. http://www.oxfordreference.com/view/10.1093/acref/9780199202720.001.0001/acref-9780199202720

Papadopoulou, D., Peristeri, E., Plemenou, E., Marinis, T. & Tsimpli, I. M. (2015). Pronoun ambiguity resolution in Greek: Evidence from monolingual adults and children. *Lingua,* 155, 98-120.

Paradis, M. (2004). *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins.

Paredes Silva, V. L. (1993). Subject omission and functional compensation: Evidence from written Brazilian Portuguese. *Language variation and change*, 5, 35-49.

Parodi, T. & Tsimpli, I. M. (2005). 'Real' and apparent optionality in second language grammars: finiteness and pronouns in null operator structures. *Second Language Research*, 21(3), 250-285.

Pascual y Cabo, D. & Rothman, J. (2012). The (Il)Logical Problem of Heritage Speaker Bilingualism and Incomplete Acquisition. *Applied Linguistics*, 1-7.

Pauwels, A. (2016). *Language maintenance and shift*. Cambridge: Cambridge University Press.

Pavlenko, A. (2014). *The bilingual mind*. Cambridge, UK: Cambridge University Press.

Pelc, L. (2001). *L1 lexical, morphological and morphosyntactic attrition in Greek-English bilinguals*. PhD dissertation, The City University of New York (CUNY).

Pérez-Leroux, A. T. & Glass, W. R. (1997). OPC effects on the L2 acquisition of Spanish. In A. T. Pérez-Leroux & W. R. Glass (Eds.), *Contemporary perspectives on the acquisition of Spanish, Vol. 1: Developing grammars* (pp. 149-165). Somerville, MA: Cascadilla Press.

Pérez-Leroux, A. T. & Glass, W. R. (1999). Null anaphora in Spanish second language acquisition: Probabilistic versus generative approaches. *Second Language Research,* 15(2), 220-49.

Perlmutter, D. M. (1978). Impersonal Passives and the Unaccusative Hypothesis. In *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society 38* (pp. 157-189). Berkeley Linguistic Society, University of California, Berkeley.

Peskova, A. (2013). Experimenting with *Pro-drop* in Spanish. *SKY Journal of Linguistics*, 26, 117-149.

Pinto, M. (2012). Delay in the Acquisition of Subject Pronouns. In S. Ferré, P. Prévost, L. Tuller & R Zebib (Eds.), *Selected Proceedings of the Romance Turn IV Workshop on the Acquisition of Romance Languages* (pp. 284-305). Newcastle: Cambridge Scholars Publishing.

Pinto, M. (2013). Tracking reference with null subjects. *Linguistics in the Netherlands*, 30, 131-145.

Pinto, M. (2014). Effects of L2 Exposure on L1 Grammar. In C. Y. Chu, C. E. Coughlin, B. Lopez Prego, U. Minai & A. Tremblay (Eds.), *Selected Proceedings of the 5th Generative Approaches to Language Acquisition North America* (GALANA 2012) (pp. 79-89). Somerville, MA: Cascadilla Proceedings Project.

Polinsky, M. (2008). Heritage Language Narratives. In D. M. Brinton, O. Kagan & S. Bauckus (Eds.), *Heritage Language Education: A New Field Emerging* (pp. 149-164). Routledge.

Polinsky, M. (2018). *Heritage languages and their speakers*. Cambridge, UK: Cambridge University Press.

Polinsky, M. & Kagan, O. (2007). Heritage languages: In the 'wild' and in the classroom. *Language and Linguistics Compass*, 1(5), 368-395.

Prada Pérez, A. (2009). Variation in subject expression in Western Romance. In S. Colina, A. Olarrea & Ana Maria Carvalho (Eds.), *Romance Linguistics 2009: Selected papers from the 39th Linguistic Symposium on Romance Languages* (LSRL) (pp. 267-284). Tucson, Arizona.

Prada Pérez, A. (2010). Subject Position in Spanish in Contact with Catalan: Language Similarity vs. Interface Vulnerability. In M. Iverson et al. (Eds.), *Proceedings of the 2009 Mind/Context Divide Workshop* (pp. 104-115). Somerville, MA: Cascadilla Proceedings Project.

Prada Pérez, A. (2018). Theoretical implications of research on bilingual subject production: The Vulnerability Hypothesis. *International Journal of Bilingualism*, 1-25.

Prentza, A. & Tsimpli, I. M. (2012). Resolution of pronominal ambiguity in Greek: syntax and pragmatics. *Μελέτες για την Ελληνική Γλώσσα, Πρακτικά της 33ης Ετήσιας Συνάντησης του Τομέα Γλωσσολογίας του Τμήματος Φιλολογίας της Φιλοσοφικής Σχολής του Α.Π.Θ.*, 197-208.

Real Academia Española & Asociación de Academias de la Lengua Española. (2011). *Nueva gramática básica de la lengua española*. Barcelona: Espasa Libros.

Real Academia Española & Asociación de Academias de la Lengua Española. (2016). *Diccionario de la lengua española*. http://dle.rae.es/?w=diccionario&origen=REDLE

Real Academia Española & Asociación de Academias de la Lengua Española. (2018). *Diccionario panhispánico de dudas*. http://www.rae.es/recursos/diccionarios/dpd

Rizzi, L. (1982). *Issues in Italian Syntax*. Dordrecht: Foris Publications.

Rizzi, L. (1986). Null Objects in Italian and the Theory of *pro*, *Linguistic Inquiry*, 17, 501-557.

Rizzi, L. (2000). Remarks on early null subjects. In M. A. Friedeman & L. Rizzi (Eds.), *The acquisition of syntax* (pp. 269-291). London: Pearson Education Limited.

Roberts, I. & Holmberg, A. (2010). Introduction: parameters in minimalist theory. In T. Biberauer, A. Holmberg, I. Roberts & M. Sheehan (Eds.), *Parametric Variation: Null Subjects in Minimalist Theory* (pp. 1-57). Cambridge, UK: Cambridge University Press.

Rothman, J. (2007a). Pragmatic solution for syntactic problems: Understanding some L2 syntactic errors in terms of pragmatic deficits. In S. Baauw, F. Dirjkoningen & M. Pinto (Eds.), *Romance Languages and Linguistic Theory 2005* (pp. 299-320). Amsterdam: John Benjamins.

Rothman, J. (2007b). Sometimes they use it, sometimes they don't: An epistemological discussion of L2 morphological production and its use as a competence measurement. *Applied Linguistics*, 28/4, 609-614.

Rothman, J. (2009a). Pragmatic deficits with syntactic consequences: L2 pronominal subjects and the syntax-pragmatics interface. *Journal of Pragmatics*, 41, 951-973.

Rothman, J. (2009b). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism*, 13(2), 155-163.

Rothman, J. & Iverson, M. (2007). Input type and parameter resetting: Is naturalistic input necessary? *International Review of Applied Linguistics in Language Teaching*, 45(4), 285-319.

Rothman, J. & Pascual y Cabo, D. (2014). Generative Approaches to Spanish Second Language Acquisition. In K. Geesling (Ed.), *Handbook of Spanish Second Language Acquisition* (pp. 46-63). Malden, MA: Blackwell.

Rothman, J. & Treffers-Daller, J. (2014). A Prolegomenon to the Construct of the Native Speaker: Heritage Speaker Bilinguals are Natives Too! *Applied Linguistics*, 35(1), 93-98.

Rothman, J., Tsimpli, I. M. & Pascual y Cabo, D. (2016). Formal linguistic approaches to heritage language acquisition. Bridges for pedagogically oriented research. In D. Pascual y Cabo (Ed.), *Advances in Spanish as a Heritage Language* (pp. 13-26). Amsterdam & Philadelphia: John Benjamins.

Roussou, A. & Tsimpli, I. M. (2006). On Greek VSO again! *Journal of Linguistics*, 42, 317-354.

Sánchez-Naranjo, J. (2013). El efecto de la referencia disjunta en español: Diversas perspectivas sobre un fenómeno complejo. *Lingüística y Literatura*, 64, 13-32.

Schmid, M. S. (2011). *Language Attrition*. Cambridge, UK: Cambridge University Press.

Schmid, M. S. & Köpke. B. (2017a). The relevance of first language attrition to theories of bilingual development. *Linguistic Approaches to Bilingualism*, 7(6), 637-667.

Schmid, M. S. & Köpke. B. (2017b). When is a bilingual an attriter? - Response to the commentaries. *Linguistic Approaches to Bilingualism*, 7(6), 763-770.

Schmitz, K. & Scherger, A. L. (2017). Effects of age and education on variable but native heritage grammars: Theoretical and empirical implications for the Null Subject Parameter. In P. Guijarro-Fuentes (Ed.), *Applied Linguistics Review. Language Acquisition and Linguistic Theory: When linguistic theory meets empirical data,* (Special Issue). DOI: 10.1515/applirev-2017-0093

Scontras, G., Fuchs, Z. & Polinsky, M. (2015). *Heritage language and linguistic theory. Frontiers in Psychology*, 6, 1545, 1-20.

Seliger, H. W. (1991). Language attrition, reduced redundancy, and creativity. In H. W. Seliger & R. M. Vago (Eds.), *First language attrition* (pp. 227-240). Cambridge, UK: Cambridge University Press.

Seliger, H. W. (1996). Primary language attrition in the context of bilingualism. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 605–625). New York, NY: Academic Press.

Serratrice, L., Sorace, A. & Paoli, S. (2004). Crosslinguistic influence at the syntax-*pragmatics* interface: Subjects and objects in Italian-English bilingual and monolingual acquisition. *Bilingualism: Language and Cognition*, 7, 183-205.

Shafto, M. A. & Tyler, L K. (2014). Language in the aging brain: the network dynamics of cognitive decline and preservation. *Science*, 346(6209): 583-587.

Sharwood Smith, M. & Van Buren, P. (1991). First language attrition and the parameter setting model. In H. W. Seliger & R. M. Vago (Eds.), *First language attrition* (pp. 17-30). Cambridge, UK: Cambridge University Press.

Sheehan, M. (2006). *The EPP and null subjects in Romance*. PhD dissertation, Newcastle University.

Shin, N. L. (2014). Grammatical complexification in Spanish in New York: 3sg pronoun expression and verbal ambiguity. *Language Variation and Change*, 26, 303-330.

Shin, N. L. (2016). Acquiring constraints on morphosyntactic variation: children's Spanish subject pronoun expression. *Journal of Child Language*, 43, 914-947. doi:10.1017/S0305000915000380

Shin, N. L. (2018). Child heritage speakers' morphosyntax. Rate of acquisition and crosslinguistic influence. In K. Potowski (Ed.), *The Routledge Handbook of Spanish as a Heritage Language* (pp. 235-253). London & New York: Routledge.

Shin, N. L. & Cairns, H. S. (2009). Subject pronouns in child Spanish and Continuity of Reference. In: J. Collentine, M. García, B. Lafford, & F. M. Marín (Eds.), *Selected Proceedings of the 11th Hispanic Linguistics Symposium* (pp. 155-64). Somerville, MA: Cascadilla Proceedings Project.

Shin, N. L. & Cairns, H. S. (2012). The Development of NP Selection in School-Age Children: Reference and Spanish Subject Pronouns, *Language Acquisition*, 19(1), 3-38.

Shin, N. L. & Erker, D. (2015). Childhood Acquisition of Spanish Subject Pronouns. In A. Carvalho, R. Orozco & N. Shin (Eds.), *Subject pronoun expression in Spanish: A crossdialectal perspective* (pp. 169-190). Washington DC: Georgetown University Press.

Sifaki, E. & Sitaridou, I. (2007). EPP Revisited: Evidence from Null Subject Languages. In E. Agathopoulou, M. Dimitrakopoulou & D. Papadopoulou (Eds.), *Selected Papers in Theoretical and Applied Linguistics, Volume 1, 17 International Symposium* (pp. 188-198). English Department, Aristotle University, Thessaloniki: Monochromia.

Sidnell, J. & Enfield, N. J. (2017). Deixis and the interactional foundations of reference. In Y. Huang (Ed.), *The Oxford Handbook of Pragmatics* (pp. 217-239). Oxford: Oxford University Press.

Silva-Corvalán, C. (1991). Spanish language attrition in a contact situation with English. In H. W. Seliger & R. M. Vago (Eds), *First Language Attrition* (pp. 151-171). Cambridge, UK: Cambridge University Press.

Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles*. Oxford, UK: Clarendon.

Silva-Corvalán, C. (2014). *Bilingual Language Acquisition: Spanish and English in the First Six Years*. Cambridge, UK: Cambridge University Press.

Silva-Corvalán, C. & Enrique-Arias, A. (2017). *Sociolingüística y Pragmática Del Español: segunda edición*. Washington DC: Georgetown University Press.

Sitaridou, I. & Kaltsa, M. (2014). Contrastivity in Pontic Greek. *Lingua*, 146, 1-27.

Slabakova, R. (2008). *Meaning in the Second Language*. Mouton de Gruyter, Berlin.

Soares da Silva, H. (2006). *O parâmetro do sujeito nulo: Confronto entre o português o e espanhol*. MA dissertation, Universidade Federal do Rio de Janeiro: Faculdade de Letras.

Sorace, A. (2004). Native language attrition and developmental instability at the syntax-discourse interface: Data, interpretations and methods. *Bilingualism: Language and Cognition*, 7 (2), 1-3.

Sorace, A. (2005). Syntactic optionality at interfaces. In: L. Cornips, & K. Corrigan (Eds.), *Syntax and Variation: Reconciling the Biological and the Social* (pp. 46-111). Amsterdam: John Benjamins.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1-33.

Sorace, A. (2012). Pinning down the concept of "interface" in bilingualism: a reply to peer commentaries. *Linguistic Approaches to Bilingualism*, 2, 209-216.

Sorace, A. (2014). Multiple grammars: Old wine in old bottles. *Second Language Research*, 30(1), 71-74.

Sorace, A. (2016). Referring expressions and executive functions in bilingualism. *Linguistic Approaches to Bilingualism*, 6(5), 669-684.

Sorace, A. & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 339-368.

Sorace, A. & Serratrice, L. (2009). Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, 13(2), 195-210.

Sorace, A., Serratrice, L. Filiaci, F. & Baldo, M. (2009). Discourse conditions on subject pronoun realization: testing the linguistic intuitions of older bilingual children. *Lingua* 119, 460-477.

Spyropoulos, V. (2007). Finiteness and control in Greek. In W. Davies & S. Dubinsky (Eds.), *New horizons in the analysis of control and raising. Studies in Natural Language and Linguistic Theory 71* (pp. 159-183). The Netherlands: Springer.

Stevenson, J. R., Crawley, R. A. & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, 9(4), 519-548. DOI: 10.1080/01690969408402130

Steward, M. (1999). *The Spanish language today*. London & New York: Routledge.

Suarez, D. (2007). Second and Third Generation Heritage Language Speakers: HL Scholarship's Relevance to the Research Needs and Future Directions of TESOL. *Heritage Language Journal*, 5(1), 27-49.

Sullivan, A. (2012). Referring in discourse. In K. Allan & K. Jaszczolt (Eds.), *The Cambridge Handbook of Pragmatics* (pp. 293-303). Cambridge, UK: Cambridge University Press.

Tamis, A. M. (2006). *Los griegos de América Latina.* Athens: Ellinika Grammata.

Tamis, A. M. (2009). Greek migration and settlement in South America. In K. Loukeris, & K. Petrakis (Eds). *The roads of the Greeks* (pp. 78-161). Athens, Greece: Polaris Publishing.

Tammer, C. (2016). *Heritage and adult L2 acquisition of empty categories in a bidialectal bilingual context. Brazilian and European Portuguese in contact*. PhD dissertation, The Arctic University of Norway.

Tomasello, M. (2008). The usage-based theory of language acquisition. The usage-based theory of language acquisition. In E. Bavin & L. Naigles (Eds.), *The Cambridge Handbook of Child Language* (pp. 89-106). Cambridge, UK: Cambridge University Press.

Torrego, E. (1998). Nominative subjects and pro-drop INFL. *Syntax*, 1, 206-219.

Treffers-Daller, J. (2009) Language Dominance and Lexical Diversity: How Bilinguals and L2 Learners Differ in their Knowledge and Use of French Lexical and Functional Items. In B. Richards, M. H. Daller, D. D. Malvern, P. Meara, J. Milton & J. Treffers-Daller (Eds.), *Vocabulary Studies in First and Second Language Acquisition* (pp. 74-90). Palgrave Macmillan, London.

Trifonas, P. P. & Aravossitas, T. (2014). Introduction. In P. P. Trifonas & T. Aravossitas (Eds.), *Rethinking heritage language education* (pp. xiii-xxi). Cambridge, UK: Cambridge University Press.

Tsimpli, I. M. (2007). First language attrition from a minimalist perspective: Interface vulnerability and processing effects. In B. Köpke, M. Schmid, M. Keijzer & S. Dostert (Eds.), *Language Attrition: Theoretical Perspectives* (pp. 83-98). John Benjamins.

Tsimpli, I. M. (2011). External interfaces and the notion of 'default'. *Linguistic Approaches to Bilingualism*, 1(1), 101-103.

Tsimpli, I. M. (2014). Early, late or very late? Timing acquisition and bilingualism. *Linguistic Approaches to Bilingualism*, 4(3), 283-313. doi: 10.1075/lab.4.3.01tsi

Tsimpli, I. M. (2017a). Crosslinguistic influence is not necessarily attrition. *Linguistic Approaches to Bilingualism*, 7(6), 759-762.

Tsimpli, I. M. (2017b). Multilingual education for multilingual speakers. *Languages, Society and Policy*. Doi:10.17863/CAM.9803

Tsimpli, I. M., Andreou, M., Agathopoulou, E. & Masoura, E. (2014). Narrative production, bilingualism and working memory capacity: A study of Greek-German bilingual children. In G. Kotzoglou et al. (Eds.), *Selected Papers of the 11th International Conference on Greek Linguistics* (pp. 1730-1942). Rhodes, University of the Aegean.

Tsimpli, I. M. & Dimitrakopoulou, M. (2007). The Interpretability Hypothesis: evidence from *wh*-interrogatives in second language acquisition. *Second Language Research*, 23(2), 215-242.

Tsimpli, I. M. & Sorace, A. (2006). Differentiating interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena. In D. Bamman, T. Magnitskaia & C. Zaller (Eds.), *Proceedings of the 30th Annual Boston Conference on Language Development* (pp. 653-664). Somerville, MA: Cascadilla Press.

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, 257–277.

Valdés, G. (2001). Heritage Language Students: Profiles and Possibilities. In J. Peyton, J. Ranard & S. McGinnis (Eds.), *Heritage Languages in America: Preserving a national resource* (pp. 37-80). McHenry, IL: The Center for Applied Linguistics and Delta Systems.

Valian, V. (2015). Innateness and learnability. In E. Bavin & L. Naigles (Eds.), *The Cambridge Handbook of Child Language* (pp. 15-36). Cambridge, UK: Cambridge University Press.

Van Esbroeck, N. (2014). *La variable sujeto: ¿Qué position ocupa el español chileno dentro del espectro de la variación dialectal?* MA dissertation, University of Gent.

Wasow, T. (2015). Ambiguity Avoidance is Overrated. In S. Winkler (Ed.), *Ambiguity Language and Communication* (pp. 31-47). De Gruyter Mouton.

Westergaard, M. (2014). Linguistic variation and micro-cues in first language acquisition. *Linguistic variation*, 14(1), 26-45.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, 106, 23-79.

White, L. (2009). Language acquisition at the interfaces: some hardy perennials and new varieties. *Mind-Context Divide: Language Acquisition and Interfaces of Cognitive-Linguistic Modules*. University of Iowa.

White, L. (2011). Second Language Acquisition at the Interfaces. *Lingua*, 121, 577-590.

Winkler, S. (2015). Exploring Ambiguity and the Ambiguity Model from a Transdisciplinary Perspective. In S. Winkler (Ed.), *Ambiguity Language and Communication* (pp. 1-25). De Gruyter Mouton.

Winter-Froemel, E. & Zirker, A. (2015). Ambiguity in Speaker-Hearer-Interaction: A Parameter-Based Model of Analysis. In S. Winkler (Ed.), *Ambiguity Language and Communication* (pp. 283-339). De Gruyter Mouton.

Yılmaz, G. & Schmid, M. (2018). First language attrition and bilingualism. In: D. Miller, F. Bayram, J. Rothman & L. Serratrice (Eds.), *Bilingual Cognition and Language: The state of the science across its subfields* (pp. 225-250). Amsterdam/Philadelphia: John Benjamins.

Yang, C. (2004). Universal Grammar, statistics or both? *TRENDS in Cognitive Sciences*, 8(10), 451-456.

Yang, C. (2015). For and against frequencies. *Journal of Child Language*, 42(2), 287-293.

Zagona, K. (2002). *The Syntax of Spanish*. Cambridge, UK: Cambridge University Press.

Zorbas, A. & Nicolaides, N. (2010). *Griegos en Chile*. Santiago: Gráfica Andes Ltda.

Zombolou, K. (2011). Attrition in Greek Diaspora. In I. Ihemere (Ed.), *Language Contact and Language Shift: Grammatical and Sociolinguistic Perspectives. Studies in Language Typology 17* (pp. 1-33). Lincom Europa.

Zubizarreta, M. L. (1998). *Prosody, focus, and word order*. Cambridge, MA: MIT Press.

Zulaica-Hernández, I. & Gelormini-Lezama, C. (2017). Introduction to Reference and Anaphora in Iberian Languages. *Journal of Psycholinguist Research*, https://doi.org/10.1007/s10936-017-9489-8.

Appendix A: Consent form

Information and Consent Form

Dear Participant,

You are invited to participate in a research project related to the use of the Greek language in Chile. This project is part of the requirements of my PhD programme in the Faculty of Modern and Medieval Languages at the University of Cambridge, which I conduct under the supervision of Dr Ioanna Sitaridou.

Information:
The purpose of this research is to study the Greek language as spoken in Chile with a focus on morphosyntax. The research will be conducted on bilingual speakers of Greek and Spanish who live in Chile. The nature of the study is both qualitative and quantitative and the research method chosen for data collection relies on recorded speech samples and a questionnaire.

Procedure:
If you agree to participate in this study, you will be asked to do the following:
1. Answer a written questionnaire on basic information about your general and linguistic background.
2. Participate in a recorded interview with the researcher. You will be asked to talk about you (and your family, if relevant) regarding Greece and Chile and your language practices.
3. Give a description of simple picture stories (in the same interview).
4. Answer a questionnaire in Greek and in Spanish on the use of language focused on certain syntactic constructions.

Right to participate, say no, or withdraw:
Participation in this project is completely voluntary. You have the right to say no or to withdraw from participating in this project at any time with any or no reason.

Contact details of the researcher:
Name: Aretousa Giannakou
Department/Institution: Department of Spanish & Portuguese, Faculty of Modern & Medieval Languages, University of Cambridge
Contact Information: ag771@cam.ac.uk mobile in Chile: 96942464

I guarantee that:
1. There are no known risks to participate in this research study.
2. All data will remain anonymous. Any information obtained will remain confidential, it will be used only for academic purposes.
3. The collection and storage of data will be according to the law on Data Protection (Data Protection Act 1998).

If you agree to participate in this research study, please sign below.

Signature: _____   Date: _____

Please feel free to contact me if you would like further information about this study.

Aretousa Giannakou

Appendix B: Background questionnaire

Background questionnaire

Date: Speaker ID:

A.   General background (ethnographic information)

Name: Age:

Date of birth: Place of birth:

Place of origin in Greece: Nationality:

Mother's nationality: Father's nationality:

Place of residence in Chile: Years of residence in Chile:

Residence in foreign countries: Yes – No   Where?   How long?

Educational level (please delete as appropriate):

Primary – Secondary – Tertiary: Non-university – University

Area of studies: Profession:

Age at which you moved to Chile (if relevant):

Have you ever lived in Greece? Yes – No   If yes, for how long?

How many times have you visited Greece (since you relocated)?

How long were the stays?

Frequency of visits to Greece:

B.   Linguistic background (language acquisition trajectory and practices)

First language(s):   Dominant language(s):

Mother's language: Father's language:

Language that you used as a child/adolescent with…

    your mother:

    your father:

    your siblings:

    your grandfathers (mother's side):
    How often?

    your grandfathers (father's side):
    How often?

Where did you go to school?

What languages did you use at…

    Primary school: Place: Language(s):

    Secondary school: Place: Language(s):

    University (1): Place: Language(s):

    University (2): Place: Language(s):

What language(s) did you learn as a child/adolescent?

What language(s) did you learn as an adult?

What language(s) do you speak now in your everyday life?

What language(s) do you speak now in your professional life?

What language(s) do you speak with your husband/partner?

What language(s) do you speak with your children?

Any other languages in the household:

Have you studied L2 Greek? Yes – No If yes, how many years?   Where?

With whom do you speak Greek?

How often do you speak Greek?

How often do you use the internet to speak to people in Greece?

How often do you read texts in Greek?

How often do you listen to Greek (e.g. from other speakers or on the radio or TV)?

How often do you write texts in Greek?

Level of Greek (self-evaluation) (please delete as appropriate):

    Beginner – Intermediate – Advanced – Near native – Native

Level of Spanish (self-evaluation) (please delete as appropriate):

    Beginner – Intermediate – Advanced – Near native – Native

Appendix C: Picture sequences used as stimuli

C.1 Horse story
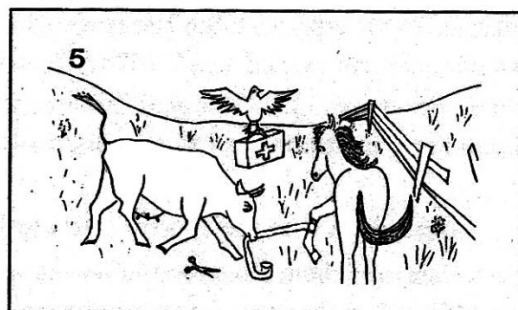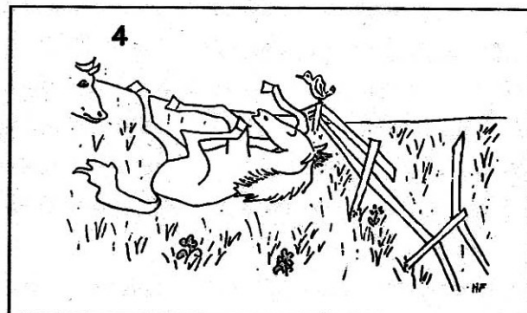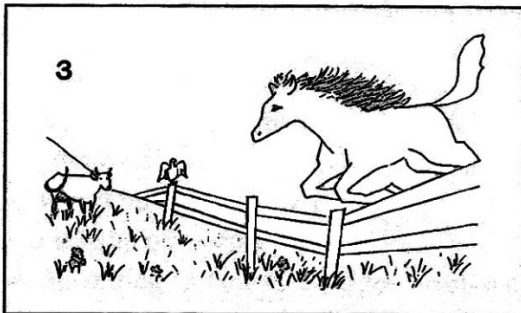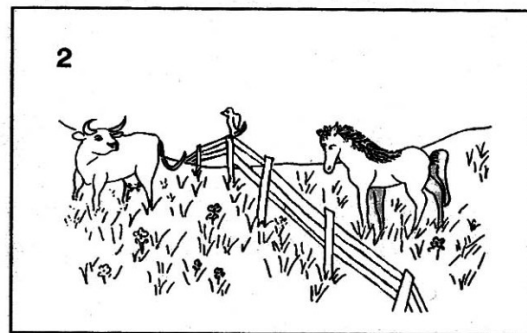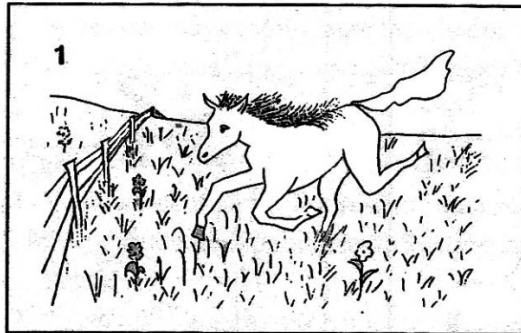
C.2 Cat story
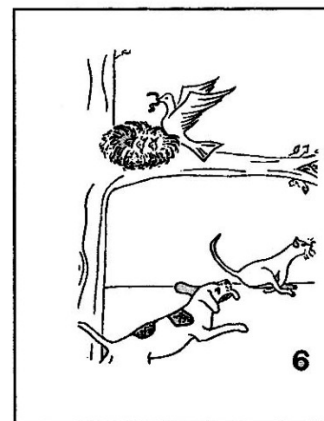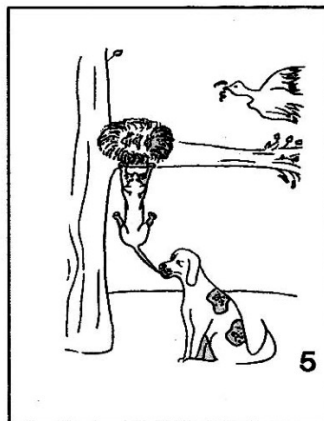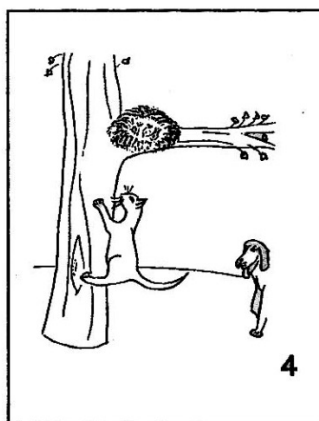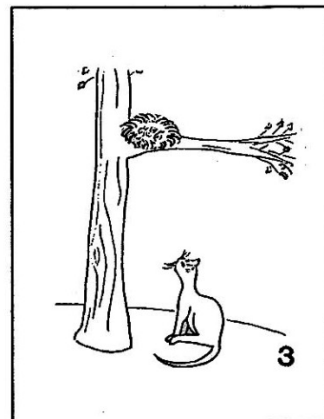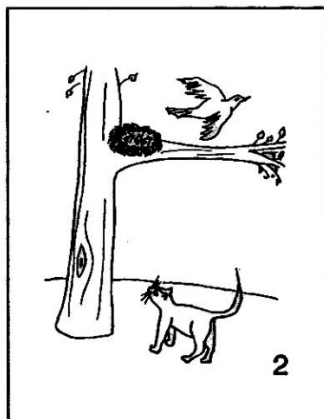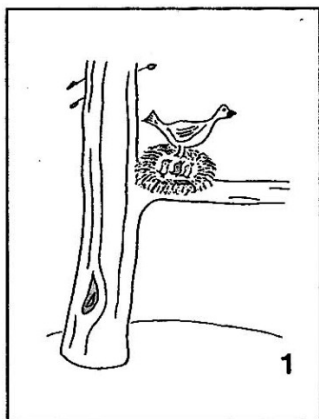
Appendix D: Anaphora resolution materials

D.1 Experimental sentences for DDN and DDO

Anaphora resolution experimental sentences with definite matrix object

		Matrix Clause	Embedded Clause	
Gender	Lang.	Definite Object	Null Subject (DDN)	Overt Subject (DDO)
Feminine	Greek	*I jaja filuse ti nosokoma*	*otan iði evaze to palto tis.*	*otan afti evaze to palto tis.*
	Spanish	*La abuela besaba a la enfermera*	*cuando ya se ponía el abrigo.*	*cuando ella se ponía el abrigo.*
	English	'The old lady was kissing the nurse	when [adverb] was putting on her coat.'	when **she** was putting on her coat.'
	Greek	*I kathijitria xeretise ti mathitria*	*otan iði pernuse to ðromo.*	*otan afti pernuse to ðromo.*
	Spanish	*La profesora saludó a la alumna*	*cuando ya cruzaba la calle.*	*cuando ella cruzaba la calle.*
	English	'The teacher greeted the student	when [adverb] was crossing the street.'	when **she** was crossing the street.'
Masculine	Greek	*O ðiefthintis xeretuse ton jatro*	*otan iði evjene apo to asanser.*	*otan aftos evjene apo to asanser.*
	Spanish	*El director saludaba al doctor*	*cuando ya salía del ascensor.*	*cuando él salía del ascensor.*
	English	'The director was greeting the doctor	when [adverb] was exiting from the lift.'	when **he** was exiting from the lift.'
	Greek	*O astinomikos iðe ton klefti*	*otan iði estrive sti yonia.*	*otan aftos estrive sti yonia.*
	Spanish	*El policía vio al ladrón*	*cuando ya doblaba la esquina.*	*cuando él doblaba la esquina.*
	English	'The policeman saw the thief	when [adverb] was turning the corner.'	when **he** was turning the corner.'

D.2 Experimental sentences for DIN and DIO

Anaphora resolution experimental sentences with indefinite matrix object

		Matrix Clause	Embedded Clause	
Gender	Lang.	Indefinite Object	Null Subject (DIN)	Overt Subject (DIO)
Feminine	Greek	*I jaja filuse mia nosokoma*	*otan iði evaze to palto tis.*	*otan afti evaze to palto tis.*
	Spanish	*La abuela besaba a una enfermera*	*cuando ya se ponía el abrigo.*	*cuando ella se ponía el abrigo.*
	English	'The old lady was kissing a nurse	when [adverb] was putting on her coat.'	when **she** was putting on her coat.'
	Greek	*I kathijitria xeretise mia mathitria*	*otan iði pernuse to ðromo.*	*otan afti pernuse to ðromo.*
	Spanish	*La profesora saludó a una alumna*	*cuando ya cruzaba la calle.*	*cuando ella cruzaba la calle.*
	English	'The teacher greeted a student	when [adverb] was crossing the street.'	when **she** was crossing the street.'
Masculine	Greek	*O ðiefthintis xeretuse enan jatro*	*otan iði evjene apo to asanser.*	*otan aftos evjene apo to asanser.*
	Spanish	*El director saludaba a un doctor*	*cuando ya salía del ascensor.*	*cuando él salía del ascensor.*
	English	'The director was greeting a doctor	when [adverb] was exiting from the lift.'	when **he** was exiting from the lift.'
	Greek	*O astinomikos iðe enan klefti*	*otan iði estrive sti yonia.*	*otan aftos estrive sti yonia.*
	Spanish	*El policía vio a un ladrón*	*cuando ya doblaba la esquina.*	*cuando él doblaba la esquina.*
	English	'The policeman saw a thief	when [adverb] was turning the corner.'	when **he** was turning the corner.'

D.3 Practice items and fillers: Greek version

Practice items:

1. Ο σκύλος πρόσεχε τα πρόβατα, ενώ η γιαγιά έφτανε στο χωριό.
   Τι πρόσεχε ο σκύλος;

2. Στο ποτάμι ο βάτραχος είδε την αλεπού που κυνηγούσε το λαγό.
   Ποιος είδε την αλεπού;

Fillers:

1. Το κορίτσι έκλαιγε, γιατί ο σκύλος έσκισε τα παπούτσια στο δωμάτιο.
   Τι έσκισε ο σκύλος;

2. Όταν κυνηγούσε η αλεπού το λαγό, την χτύπησε ένα αυτοκίνητο.
   Ποιόν χτύπησε το αυτοκίνητο;

3. Η τίγρης είδε τη καμηλοπάρδαλη και το ελάφι κοντά στη λίμνη.
   Ποιο ζώο είναι πιο ψηλό;

4. Ο δικηγόρος κρατούσε σημειώσεις στον υπολογιστή, ενώ βρισκόταν στη βιβλιοθήκη.
   Πού βρισκόταν ο υπολογιστής;

5. Ο νεαρός έδωσε τα λουλούδια στην κοπέλα, ενώ άκουγε μουσική.
   Ποιος έδωσε τα λουλούδια;

6. Η γάτα κρύφτηκε τρομαγμένη γιατί είδε μια τίγρη στη τηλεόραση.
   Τι τρόμαξε τη γάτα;

7. Όταν έπεσε η μπάλα από τη πολυθρόνα έσπασε ο καθρέφτης.
   Πού ήταν η μπάλα;

8. Ο σκύλος γάβγιζε, όταν η κοπέλα διάβαζε το μάθημά της στο πιάνο.
   Τι ενοχλούσε το κορίτσι;

9. Ενώ ο ανεμιστήρας δούλευε, η γάτα ανέβηκε στον υπολογιστή του γραφείου.
   Τι ανέβηκε στον υπολογιστή;

10. Η μαθήτρια φωτογράφισε την αρκούδα και μετά έψαχνε το κλουβί του ελέφαντα.
    Τι φωτογράφισε η μαθήτρια;

11. Η αρκούδα κυνήγησε το γορίλλα ενώ ο ελέφαντας έτρωγε στο κλουβί του.
    Τι κυνήγησε η αρκούδα;

12.   Η αθλήτρια έβλεπε τους αγώνες στην τηλεόραση, αφού έκλεισε το ραδιόφωνο.
        Τι έβλεπε η αθλήτρια;

13.   Το παγωτό που έτρωγε το παιδί έσταξε και λέρωσε το παντελόνι του.
        Τι λέρωσε το παντελόνι του;

14.   Ο καναπές και το ραδιόφωνο είναι στο δωμάτιο με τον καθρέφτη.
        Από πού ακούγεται μουσική;

15.   Ο ανεμιστήρας δρόσιζε τη γάτα, ενώ αυτή έπαιζε με τη μπάλα.
        Με τι έπαιζε η γάτα;

16.   Ο σκύλος όρμησε στη γάτα που ήθελε να φάει το καναρίνι.
        Τι ήθελε να φάει η γάτα;

D.4 Practice items and fillers: Spanish version

Practice items:

1.   El perro cuidaba las ovejas mientras la abuela llebaga al pueblo.
        ¿Qué cuidaba el perro?

2.   En el río la rana vio al zorro que cazaba el conejo.
        ¿Quién vio al zorro?

Fillers:

1.   La niña lloraba porque el perro le rompió los zapatos en la pieza.
        ¿Qué rompió el perro?

2.   Cuando la gata perseguía al conejo la atropelló un auto.
        ¿A quién atropelló el auto?

3.   El tigre vio a la jirafa y el ciervo cerca del lago.
        ¿Qué animal es más alto?

4.   El abogado tomaba notas al computador mientras se encontraba en la biblioteca.
        ¿Dónde se encontraba el computador?

5.   El joven le dio las flores a la chica mientras escuchaba música.
        ¿Quién dio las flores?

6.   El gato se escondió asustado porque vio a un tigre en la tele.
        ¿Qué asustó al gato?

7.   Cuando se cayó la pelota del sillón se rompió el espejo.
        ¿Dónde estaba la pelota?

8.   El perro ladraba cuando la chica estudiaba su lección de piano.
        ¿Qué molestaba a la chica?

9.   Mientras el ventilador estaba funcionando el gato subió al computador de la oficina.
        ¿Quién subió al computador?

10.   La alumna fotografió al oso y luego buscó la jaula del elefante.
        ¿Qué fotografió la alumna?

11.   El oso cazó al gorila mientras el elefante comía en su jaula.
        ¿A quién cazó el oso?

12.   La atleta veía los juegos en la tele después de apagar la radio.
        ¿Qué veía la atleta?

13.   El helado que comía el chico se derritió y le manchó el pantalón.
        ¿Qué manchó el pantalón?

14.   El sofá y la radio están en la pieza con el espejo.
        ¿De dónde se escucha música?

15.   El ventilador refrescaba al gato mientras él jugaba con la pelota.
        ¿Con qué jugaba el gato?

16.   El perro atacó al gato que quería comer al canario.
        ¿Qué quería comer el gato?

D.5 Items for demonstrative pronoun resolution

1. La abuela besaba a la enfermera, cuando ésta se ponía el abrigo.
   ¿Quién se ponía el abrigo?

2. La abuela besaba a una enfermera, cuando ésta se ponía el abrigo.
   ¿Quién se ponía el abrigo?

3. La profesora saludó a la alumna, cuando ésta cruzaba la calle.
   ¿Quién cruzaba la calle?

4. La profesora saludó a una alumna, cuando ésta cruzaba la calle.
   ¿Quién cruzaba la calle?

5. El policía vio al ladrón, cuando éste doblaba la esquina.
   ¿Quién doblaba la esquina?

6. El policía vio a un ladrón, cuando éste doblaba la esquina.
   ¿Quién doblaba la esquina?

7. El director saludaba al doctor, cuando éste salía del ascensor.
   ¿Quién salía del ascensor?

8. El director saludaba al doctor, cuando éste salía del ascensor.
   ¿Quién salía del ascensor?

Appendix E: Participants

E.1 Monolinguals

<table>
<tr><th colspan="7" align="center">Monolinguals</th></tr>
<tr><th colspan="5" align="center">Spanish monolinguals</th><th colspan="2" align="center">Greek monolinguals</th></tr>
<tr><th colspan="3" align="center">Main study</th><th colspan="2" align="center">Demonstratives (AR)</th><th></th><th></th></tr>
<tr><th>#</th><th>Speaker ID</th><th>Age</th><th>Speaker ID</th><th>Age</th><th>Speaker ID</th><th>Age</th></tr>
<tr><td>1</td><td>S65</td><td>70</td><td>S164</td><td>28</td><td>S91</td><td>38</td></tr>
<tr><td>2</td><td>S66</td><td>29</td><td>S165</td><td>44</td><td>S94</td><td>66</td></tr>
<tr><td>3</td><td>S67</td><td>59</td><td>S166</td><td>55</td><td>S95</td><td>48</td></tr>
<tr><td>4</td><td>S68</td><td>28</td><td>S167</td><td>48</td><td>S96</td><td>46</td></tr>
<tr><td>5</td><td>S75</td><td>56</td><td>S168</td><td>37</td><td>S97</td><td>38</td></tr>
<tr><td>6</td><td>S76</td><td>44</td><td>S170</td><td>38</td><td>S99</td><td>36</td></tr>
<tr><td>7</td><td>S86</td><td>77</td><td>S171</td><td>41</td><td>S103</td><td>64</td></tr>
<tr><td>8</td><td>S87</td><td>40</td><td>S172</td><td>31</td><td>S105</td><td>70</td></tr>
<tr><td>9</td><td>S69 (only narratives)</td><td>70</td><td>S173</td><td>48</td><td>S107</td><td>33</td></tr>
<tr><td>10</td><td>S70 (only narratives)</td><td>39</td><td>S174</td><td>32</td><td>S88 (only narratives)</td><td>34</td></tr>
<tr><td>11</td><td>S71 (only narratives)</td><td>39</td><td>S175</td><td>62</td><td>S89 (only narratives)</td><td>36</td></tr>
<tr><td>12</td><td>S72 (only narratives)</td><td>40</td><td>S176</td><td>85</td><td>S90 (only narratives)</td><td>46</td></tr>
<tr><td>13</td><td>S73 (only narratives)</td><td>38</td><td>S177</td><td>28</td><td>S92 (only narratives)</td><td>31</td></tr>
<tr><td>14</td><td>S77 (only narratives)</td><td>37</td><td>S178</td><td>60</td><td>S93 (only narratives)</td><td>52</td></tr>
<tr><td>15</td><td>S78 (only narratives)</td><td>45</td><td>S179</td><td>29</td><td>S98 (only narratives)</td><td>43</td></tr>
<tr><td>16</td><td>S79 (only narratives)</td><td>39</td><td>S180</td><td>60</td><td>S100 (only narratives)</td><td>48</td></tr>
<tr><td>17</td><td>S82 (only narratives)</td><td>39</td><td>S181</td><td>30</td><td>S101 (only narratives)</td><td>42</td></tr>
<tr><td>18</td><td>S83 (only narratives)</td><td>39</td><td>S182</td><td>27</td><td>S102 (only narratives)</td><td>29</td></tr>
<tr><td>19</td><td>S84 (only narratives)</td><td>39</td><td>S183</td><td>48</td><td>S104 (only narratives)</td><td>68</td></tr>
<tr><td>20</td><td>S85 (only narratives)</td><td>37</td><td>S184</td><td>73</td><td>S106 (only narratives)</td><td>41</td></tr>
<tr><td>21</td><td>S117 (only AR)</td><td>34</td><td></td><td></td><td>S108 (only AR)</td><td>80</td></tr>
<tr><td>22</td><td>S118 (only AR)</td><td>20</td><td></td><td></td><td>S109 (only AR)</td><td>38</td></tr>
<tr><td>23</td><td>S119 (only AR)</td><td>20</td><td></td><td></td><td>S110 (only AR)</td><td>58</td></tr>
<tr><td>24</td><td>S120 (only AR)</td><td>71</td><td></td><td></td><td>S111 (only AR)</td><td>79</td></tr>
<tr><td>25</td><td>S121 (only AR)</td><td>42</td><td></td><td></td><td>S112 (only AR)</td><td>16</td></tr>
<tr><td>26</td><td>S122 (only AR)</td><td>20</td><td></td><td></td><td>S113 (only AR)</td><td>16</td></tr>
<tr><td>27</td><td>S123 (only AR)</td><td>38</td><td></td><td></td><td>S114 (only AR)</td><td>60</td></tr>
<tr><td>28</td><td>S124 (only AR)</td><td>62</td><td></td><td></td><td>S115 (only AR)</td><td>39</td></tr>
<tr><td>29</td><td>S125 (only AR)</td><td>60</td><td></td><td></td><td>S135 (only AR)</td><td>21</td></tr>
<tr><td>30</td><td>S126 (only AR)</td><td>48</td><td></td><td></td><td>S137 (only AR)</td><td>73</td></tr>
<tr><td>31</td><td>S127 (only AR)</td><td>32</td><td></td><td></td><td>S144 (only AR)</td><td>44</td></tr>
<tr><td>32</td><td>S132 (only AR)</td><td>85</td><td></td><td></td><td></td><td></td></tr>
</table>

E.2 Bilinguals

Immigrants

#	Speaker ID	Age	AoB	LoR	YoB	Greek Exposure
1	S60	26	24	1	2	daily
2	S133 (only AR)	37	25	1	12	daily
3	S10	28	25	3	3	regular
4	S147 (only AR)	31	28	3	3	regular
5	S59	45	41	3	3	daily
6	S21	37	32	3	5	daily
7	S143	22	19	3	3	occasional
8	S12 (only narratives)	29	25	4	4	regular
9	S20	39	35	4	4	regular
10	S46 (only narratives)	36	28	4	8	daily
11	S52	39	35	4	4	daily
12	S05 (only narratives)	37	32	5	5	daily
13	S34	34	28	5	6	daily
14	S55	35	30	5	5	daily
15	S63 (only narratives)	29	19	5	10	regular
16	S40	38	32	6	6	regular
17	S01	44	31	6	13	regular
18	S53	35	24	7	11	daily
19	S22 (only narratives)	35	27	8	9	daily
20	S61	39	29	10	10	daily
21	S03	36	24	12	12	daily
22	S140 (only AR)	40	26	14	14	regular
23	S11	52	36	16	16	regular
24	S142	34	13	17	21	occasional
25	S19	35	13	19	22	regular
26	S18	38	6	20	32	daily
27	S32	84	42	29	42	daily
28	S37	74	23	51	51	occasional
29	S09	80	26	54	54	daily
30	S02	84	30	56	56	rare
31	S149	81	24	58	58	occasional
32	S54	81	22	59	59	rare
33	S43	84	24	60	60	occasional
34	S08	82	19	63	63	daily
35	S51	76	13	63	63	occasional
36	S162 (only AR)	72	8	64	64	occasional
37	S56 (only narratives)	81	19	65	65	occasional
38	S58	77	11	66	66	occasional
39	S146	81	15	66	66	rare
40	S157 (only AR)	87	21	66	66	rare
41	S44 (only biographical interview)	89	15	74	74	rare

Heritage speakers

#	Speaker ID	Age	AoB	YoB	Greek Exposure	Greek heritage	Proficiency		
							CEFR	Grammaticality score	Proficiency
1	S30	49	0	49	regular	25% (grandfather)	A1/A2	66%	Basic
2	S153	57	0	57	occasional	100% (both parents)	A2/B1	74%	Intermediate
3	S50	43	0	43	rare	25% (grandfather)	A2/B1	76%	Intermediate
4	S156	63	0	63	occasional	100% (both parents)	B1	76%	Intermediate
5	S23	29	0	29	regular	75% (father HS & mother)	B1	78%	Intermediate
6	S39	63	0	63	regular	50% (both sets of grandparents)	B1	80%	Intermediate
7	S159	62	0	62	occasional	100% (both parents)	B1	80%	Intermediate
8	S148	63	0	63	occasional	100% (both parents)	B1/B2	82%	Intermediate
9	S150	54	0	54	occasional	100% (both parents)	B1	82%	Intermediate
10	S36	57	0	57	occasional	25% (one set of grandparents)	C1	88%	Advanced
11	S160	65	0	65	occasional	100% (both parents)	B2/C1	88%	Advanced
12	S154	59	0	59	occasional	100% (both parents)	B2/C1	92%	Advanced
13	S42	57	0	57	daily	100% (both parents)	C1	92%	Advanced
14	S155	67	0	67	occasional	100% (both parents)	B2/C1	92%	Advanced
15	S152	57	0	57	regular	100% (both parents)	B2/C1	94%	Advanced
16	S139	30	0	30	regular	50% (father)	B2	94%	Advanced
17	S26	16	0	16	daily	50% (mother HS & grandmother)	B2/C1	96%	Near-native
18	S31	44	0	44	regular	100% (both parents)	B2/C1	96%	Near-native
19	S48	43	0	43	occasional	100% (both parents)	C1	98%	Near-native
20	S25	48	0	48	daily	100% (both parents)	C2	100%	Near-native
21	S45	65	0	65	occasional	100% (both parents)	C1	100%	Near-native
22	S161 (only AR)	85	0	85	occasional	100% (both parents)	A2	-	Basic
23	S145 (only AR)	72	0	72	occasional	50% (father)	B1	-	Intermediate

L2 speakers

#	Speaker ID	Age	AoB	YoB	Greek Exposure	Greek heritage/connection	Proficiency		
							CEFR	Grammaticality score	Proficiency
1	S28 (only narratives)	50	47	3	occasional	25% (grandfather)	A1	62%	Basic
2	S64	31	26	5	regular	Consort	A1/A2	68%	Basic
3	S29	69	62	7	occasional	50% (father)	A1	70%	Basic
4	S57 (only narratives)	79	20	59	regular	Consort	A1	72%	Basic
5	S151	29	27	2	regular	25% (grandfather)	B1	78%	Intermediate
6	S141	24	22	2	regular	25% (grandmother)	B1	78%	Intermediate
7	S47	30	15	15	regular	25% (grandfather)	B1	78%	Intermediate
8	S24	30	26	4	occasional	25% (one set of grandparents)	B1	78%	Intermediate
9	S38	51	40	11	occasional	Migration to Greece (4 years)	B1	80%	Intermediate
10	S138	68	25	43	regular	Consort	B2/C1	86%	Advanced
11	S27	30	15	15	occasional	25% (grandfather)	B2	88%	Advanced
12	S14	27	20	7	occasional	25% (grandfather)	B2	90%	Advanced
13	S49	25	21	4	occasional	25% (grandfather)	B2	90%	Advanced
14	S16	36	19	17	regular	Migration to Greece (8 years)	B2	94%	Advanced
15	S15	58	17	41	regular	Migration to Greece (8 years)	C2	96%	Near-native
16	S33	43	26	17	regular	25% (grandfather)	C2	96%	Near-native
17	S134	37	26	11	daily	Consort	C2	96%	Near-native
18	S17	47	8	39	daily	Migration to Greece (25 years)	C2	98%	Near-native
19	S13	71	20	51	regular	25% (grandfather)	C2	98%	Near-native
20	S35 (only narratives)	54	25	29	daily	Migration to Greece (24 years)	C2	100%	Near-native
21	S158 (only AR)	32	23	9	regular	25% (grandfather)	A2/B1	-	Intermediate