FACULTÉ
DE LETTRES, TRADUCTION
& COMMUNICATION
UNIVERSITÉ LIBRE DE BRUXELLES

Bilingual implications

Using code-switching data to inform linguistic theory

Thesis submitted by Emma VANDEN WYNGAERD

with a view to obtaining the PhD Degree in Languages, Humanities and Traductology ("Docteur en Lettres, Traduction et Communication")

Academic year 2020-2021

Supervisor: Philippe DE BRABANTER

LADISCO
Center for Research in Linguistics

BILINGUAL IMPLICATIONS

EMMA VANDEN WYNGAERD

Using code-switching data to inform linguistic theory

Truth is much too complicated to allow anything but approximations.

— John von Neumann, 1947

Everything is fine, the world is still on the turtle.

— John Green, 2020

voor mijn lievekes

ABSTRACT

In the last few decades, there has been increased interest in the incorporation of data from bi- and multilingual individuals in linguistic theory: from second language acquisition and language attrition to heritage varieties and code-switching. This dissertation discusses a range of ways in which code-switching data can provide insight into the mechanisms that underlie linguistic structures. The data will be analysed within the framework of Minimalist Generative syntax and Distributed Morphology.

The first part investigates grammatical gender assignment in code-switching between English, a language without grammatical gender, and two languages *with* grammatical gender: French and Belgian Dutch. These languages have comparable, but different gender systems. French has two genders: masculine and feminine, whereas Belgian Dutch adds a third: neuter. The study in this part of the dissertation compares gender assignment strategies in bilinguals with different profiles. In addition, the code-switching data provide evidence against the default status of neuter in Belgian Dutch.

The second part focuses on word order and includes two studies: one on verb-second word order in Dutch-English code-switching and one on adverb placement in English-French and Dutch-English code-switching. The verb-second chapter identifies a lacuna in the traditional Generative analysis for verb second and uses the CS data to address this. The chapter on adverb position looks at placement of the adverb between the verb and its direct object, which is allowed in Dutch and French, but not in English. For all domains investigated, it is found that the finite verb predicts word order.

Taken together, these studies demonstrate that bilingual data can shine a light on elements of the theory of grammar which remain in the shadows when only monolingual data is used.

RÉSUMÉ

Les dernières décennies ont vu croître l'intérêt pour l'intégration à la réflexion en linguistique théorique des données produites par des locuteurs/trices bilingues ou multilingues, que celles-ci concernent l'acquisition d'une langue seconde, l'attrition, les langues d'héritage ou l'alternance codique. Le présent travail développe plusieurs exemples où les données issues de l'alternance codique éclairent les mécanismes qui sous-tendent les structures linguistiques. Les données recueillies sont interprétées dans le cadre de la syntaxe générative minimaliste et de la morphologie distribuée (« distributed morphology »).

Dans un premier temps, nous analysons l'attribution du genre grammatical dans l'alternance entre l'anglais, d'une part, et le français et le néerlandais de Belgique, de l'autre. Alors qu'il n'y a pas en anglais de genre grammatical, le français et le néerlandais de Belgique marquent ce genre, mais de façon différente : si le français distingue deux genres, masculin et féminin, le néerlandais de Belgique y adjoint un troisième, le neutre. Dans cette partie de la thèse, nous dressons le profil des stratégies d'attribution du genre auprès de deux types distincts de bilingues et nous établissons également que le neutre n'est pas le genre par défaut en néerlandais de Belgique.

Dans un second temps, nous nous penchons sur l'ordre des constituants. Dans une première étude, nous examinons l'ordre des mots avec « verbe second » (V2) dans l'alternance anglais-néerlandais. Nous abordons ensuite le placement de l'adverbe dans l'alternance anglais-français et anglais-néerlandais. Le chapitre consacré à V2 identifie une lacune dans la littérature générative et tire profit des données de l'alternance pour y proposer une solution. Le chapitre consacré à l'adverbe s'intéresse au placement de celui-ci entre le verbe et son objet, position licite en français et néerlandais mais pas en anglais. Dans ces deux études, il apparaît que c'est la langue du verbe à la forme finie qui prédit l'ordre des constituants.

L'ensemble des recherches ici réunies démontre que les données bilingues mettent en lumière des aspects de la théorie grammaticale qui restent dans l'ombre lorsque le chercheur se limite à des données monolingues.

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

Vanden Wyngaerd, Emma (2017). "The adjective in Dutch-French codeswitching: Word order and agreement." In: *The International Journal of Bilingualism* 21.4, pp. 454–473.

Vanden Wyngaerd, Emma (2018). "Adapting the traditional account of V2 using bilingual data." In: *Papers of the LSB*. Vol. 12.

Vanden Wyngaerd, Emma (2020). "C^0 and Dutch-English codeswitching." In: *Ampersand* 7.100060, pp. 1–22.

ACKNOWLEDGMENTS

My love for linguistics started at Leiden University. I was introduced to so many interesting topics and got to know so many great people. Thanks to Sofie van der Pol for the weekly dinners and Suzanne van der Meer and others in the MPhil room for commiserations. I am also very thankful to Johan, Drea, Dries and Oriane, who provided me with a home away from home in Leiden which, at times, I sorely needed.

I am indebted to Dr Parafita Couto for putting me on the code-switching path when she supervised my master-thesis. Thank you Maricarmen, for helping me write an interesting and engaging MA-thesis, encouraging me to submit that work to conferences and journals, and involving me in your projects. I owe many of my connections in the CS world to you.

The next step took me to the Université libre de Bruxelles. Thanks to my fellow PhD students at LaDisco and the L&L department. I loved getting to talk with you about our research and other things, listening to your talks, … Thanks to my office- and hall-mates Gregory Watson, Astrid De Wit, Evie Tops and Jürgen Jaspers for fun distractions during the workday, drinks, karaoke and cake. I would also like to thank my students who have taught me so much, though surely it is supposed to be the other way round.

I would like to thank my supervisor, Dr De Brabanter for his guidance throughout the years, useful and detailed comments on draft upon draft of (parts of) this dissertation and for stocking the office with ever fancier biscuits and chocolates.

I am indebted to all the participants who took part in my surveys. Without them, this dissertation wouldn't have seen the light of day.

Six years is a long time, even by the standards of academic publishing. Some parts of this dissertation have previously been submitted to scientific journals. Because of the feedback I received from reviewers, the relevant parts of this work were substantially improved. Similarly, the members of my dissertation committees, Dr Geelhand, Dr Jaspers, Dr Kissine, Dr Liceras and Dr Parafita Couto, provided me with valuable input. I am thankful to them and the reviewers for generously donating their time to improve my work.

Since I started my linguistic adventure twelve years ago, my life has changed a lot. I moved countries (twice!), got married, built a house, we grew our family. Luckily, some things also stayed the same. I am grateful beyond words to my parents Katrin and Guido, sisters Anna and Linnea, husband Brian, daughter Ina and her sibling-to-be. They support/cheer/distract/hold/laugh with/love/listen to/comfort/help me in so many different ways, not only during this project, but also in the years before and the years to come.

Ik zie jullie graag.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS AND ABBREVIATIONS

2AFC	two-alternative forced-choice task
AC	Adjacency Condition
AIC	Akaike's Information Criterion
AJT	acceptability judgment task
ANOVA	analysis of variance
AoA	age of aquisition
BD	Belgian Dutch
BIC	Bayesian Information Criterion
CI	confidence interval
CP	complementiser phrase
CS	code-switching
DM	Distributed Morphology
DP	determiner phrase
EFM	Exoskeletal Frame Model
EL	Embedded Language
FI	Feature Inheritance
FinP	finiteness phrase
FocP	focus phrase
G&B	Government and Binding
L_1	first language
L_2	second language
LF	Logical Form
LOLI	lone other-language item

ML Matrix Language
MLF Matrix Language Frame
MP Minimalist Program
MOP Morpheme Order Principle
OT Optimality Theory
PF Phonological Form
PFIC PF Interface Condition
StD standard deviation
SD Standard Dutch
SE standard error
SMP System Morpheme Principle
SOV subject object verb
SVO subject verb object
TAG Tree-Adjoining Grammar
TE translational equivalent
TP tense phrase
USP Uniform Structure Principle
V2 verb second
VI Vocabulary Insertion
VP verb phrase
XSM eXoSkeletal Model

GLOSSING CONVENTIONS

This dissertation follows the Leipzig Glossing Rules (Comrie et al. 2008). Abbreviations and conventions not included in those rules are listed below.

bold faced text	Dutch language material
italic text	French language material
SMALL CAPS	English language material
underlined text	Spanish
$\sqrt{}$	lexical root in the DM sense
AG	agentive
C	common gender
DIM	diminutive
INTERJ	interjection
NUM	number

INTRODUCTION

> Code-switching sentences look as if they had been generated in the manner of a kidnapper's ransom note by clipping words and phrases out of bilingual newspapers and pasting them end to end.
>
> — Woolford (1983, p 522)

Despite the first impression that many people have of code-switching – the seamless alternation between two or more languages in one conversation – over the last few decades it has become clear that there are rules that regulate how multilinguals mix their languages.

Chomsky (1965) famously said that "[l]inguistic theory is concerned primarily with an ideal speaker-listener" (p 3). Even though multilinguals far outnumber monolinguals world-wide (Tucker 1999, p 1), it seems like the ideal speaker-listener was considered to be the monolingual:

> [A]nything but a monolingual speaker is argued to be too complicated as an object of study. When attempting to discover the underlying principles of the faculty of language, we need to study "pure cases" to ensure that what we discover has not been affected by other factors.
>
> Lohndal (2013, p 216)

However, over the last two decades, the realisation has grown that rather than muddying the waters, linguistic data from bi- and multilingual individuals can actually help us see things more clearly. Studying the interaction of two (or more!) grammars can provide valuable insights into the investigation of linguistic structure. In Generative approaches to code-switching, the following question (1) has come to lead the field.

(1) A QUESTION
How do bilingual data help us understand the architecture of grammar?

This dissertation follows in that same tradition. Its subtitle provides the overarching theme of the present work. I will be using code-switching data to inform linguistic theory. I will address several research questions within this overarching theme using acceptability judgment tasks in which code-switched stimuli of two different language pairs are judged: English-French and Dutch-English.

The first set of research questions centre on grammatical gender. As we will see in chapter 5, code-switching (CS) data have been used to draw conclusions about the status of default gender in a variety of languages. Which gender is default in French is pretty uncontroversial, but which one is the default in Dutch is less clear. The data discussed in this chapter will come to bear on that question. In addition, the CS literature has identified several gender agreement strategies for CS in a variety of language pairs and the second question addressed in this chapter is whether these can be found as well for the language pairs under investigation in this dissertation.

The second set of research questions involves word order. First, I turn my attention to verb second word order. As we will see in chapter 7, there seems to be an issue with the traditional Generative account for verb second word order when it comes to incorporating CS data. This account seems to make no predictions for bilingual data. So the question addressed in this chapter is: Do the bilingual data pose a problem for the traditional analysis of verb second word order? If so, can we fine-tune this analysis to account for the CS judgment patterns found in those data?

Second, I highlight the position of adverbials. While the monolingual literature seems pretty much in agreement on what drives the difference in adverbial placement between Dutch, English and French, the recent CS literature had cast some doubt on the

common assumptions. Chapter 8 presents judgment data on the position of the adverb in English-French and Dutch-English CS with the aim of answering the following question: What determines the position of the adverb? Is it verb movement as has been assumed in the monolingual literature, or are there other factors that come into play?

LANGUAGES

This dissertation discusses new CS data collected from two language pairs: Dutch-English and English-French. The motivation is partly practical: these are three languages that I speak with varying levels of proficiency. However, the choice is not just determined by this fact.

While it is certainly not the case that CS in these language pairs has escaped investigation entirely (see for example Brasart 2013; Broersma 2011; Brown 1986; Schatz 1989), there is not a wealth of literature on the topic, especially within a Generative framework. Hence, in this very narrow sense, these language pairs provide quasi-unexplored territory. Individually, however, these three languages are extremely well described. This provides us with an excellent map to commence our exploration.

Investigations of CS between Spanish and English form the bulk of the CS literature, both Generative and otherwise. Since Spanish and French are closely related, this means that a lot of the insights gained in the field could presumably be expanded to the English-French language pair. Those insights can guide the investigation.

Dutch-English CS provides us with perhaps a different insight. This is CS between closely related languages, which nonetheless present some interesting typological differences. While the Generative literature on pairs of similar languages is not abundant, it is not inexistent either (Van Dulm 2007).

The variety of Dutch targeted in this dissertation is Belgian Dutch. The relationship between the varieties of Dutch spoken in Belgium and the Netherlands is – both diachronically and synchronically – complex. While Dutch spoken in the Netherlands is heavily standardised, the Dutch spoken in Belgium is currently undergoing standardisation (Grondelaers and Hout 2011, p 199). The spoken variety in Belgium that is closest to the standardised Dutch spoken in the Netherlands is sometimes known as VRT Dutch, after the public broadcaster of the same name. VRT Dutch is the spoken counterpart of written Standard Dutch. While it is phonologically quite different from Netherlandic spoken Standard Dutch, both varieties are (morpho)syntactically very similar (Grondelaers and Hout 2011, p 217–218). However, VRT Dutch is quite aptly named, as it is only "rarely spoken in practice" outside the broascasting context (De Caluwe 2009, p 19). "Standard Dutch" will be used in this dissertation to refer to both the written standard, and the morphosyntactic properties shared by VRT Dutch and spoken Netherlandic Standard Dutch.

If VRT Dutch is not spoken in practice, which variety of Dutch *is* then the colloquial standard in Belgium? "Tussentaal"[1] (literally: in-between language) is a widely spoken variety of Dutch in Belgium which displays a mix of characteristics of both Standard Dutch and Belgian Dutch dialects. Interestingly, "[a]lthough Tussentaal is immediately recognisable to Belgian listeners, it cannot easily be characterised in terms of necessary and sufficient features" (Grondelaers and Hout 2011, p 222). While it can vary regionally, Tussentaal is "often perceived to be especially colored by features of which the origins can be traced to the central dialect area (consisting of the provinces of Antwerp and Flemish Brabant, as opposed to the peripheral (West-Flemish and Limburg) dialect regions" (Jaspers and Van Hoof 2015, fn 10).

1 I follow the majority of the recent literature in naming this variety "Tussentaal", as I agree with De Caluwe's (2009) argument that "in-between language" is not an accurate *descriptive* label for what has become the *de facto* colloquial Dutch standard in Belgium.

Whenever "Belgian Dutch" is used in this dissertation, it refers to Tussentaal.

The difference between Standard Dutch (SD) and Belgian Dutch (BD) will prove relevant especially in part ii, as the gender systems of SD and BD diverge. While there is quite some regional variation in BD, masculine gender agreement is a reliable marker of BD (Jaspers and Van Hoof 2015, p 11–12). For the word order effects that are investigated in part iii, BD and SD have the same properties.

SOME PRELIMINARIES

"Code-switching" has different spellings, ("codeswitching", "code switching") but I will always spell it with a hyphen, the most common spelling in the literature, unless direct quotes contain a different spelling. For my definition of CS, see chapter 2.

Throughout the dissertation, in the code-switched examples and the running text, English language elements will be in SMALL CAPITALS, French will be in *italics* and Dutch will be **bold faced**. Since a lot of the literature focuses on data from ENGLISH-Spanish CS, this pair also gets its own convention and Spanish will be underlined. Whenever data from other language pairs is discussed, the typographic convention will be made explicit in the surrounding text. Language pairs are ordered alphabetically, not according to any other criterion.

In the literature on bi- and multilingualism, there are several terms that are often used and that I take for granted here. "L_1", "L_2" and "L_3" (respectively, first, second and third language) will be used to refer to the order of acquisition of the languages of the multilingual individual. Another useful term is "language dominance". Dominance has been defined in several ways, most notably in terms of proficiency, fluency or frequency of use. Multilinguals who have one language which is more dominant than the

other(s) are sometimes referred to as "unbalanced" bi/multilinguals, while "balanced" bi/multilinguals have similar command of all their languages. See Treffers-Daller (2015) for a critical review of the concept "balanced bilingual".

In the acquisition literature, the distinction is often made between sequential (or successive) and simultaneous bilinguals. The latter have two native languages, while the latter start acquiring their second language at a later stage. Sometimes used synonymously to these two terms are the terms "early" and "late" bilinguals. I will stick to the opposition between late and early bilinguals, as my cut-off point between the two groups is age six (see section 5.2.3). This means that my early bilinguals include both true simultaneous bilinguals and early sequential bilinguals.

Like this one!

Finally, on a different topic, I want to mention the sidenotes. This dissertation is perhaps the most personal piece of work that I will get to write. While I have largely stuck to an academic style in the text of this dissertation, the sidenotes provide a little bit of personal, informal commentary.

OUTLINE

This dissertation consists of three parts. Part i contains three chapters that together form the backdrop for this dissertation. While each of the sections of this part could probably form the subject of an entire monograph, I have limited myself to the notions and concepts that I will rely on in parts ii and iii. Chapter 1 sets out the theoretical framework of the thesis, chapter 2 reviews the literature on the phenomenon of CS and chapter 3 discusses some methodological considerations. Parts ii and iii form the bulk of this dissertation.

In part ii I take an in-depth look at CS and grammatical gender. Chapter 4 provides a detailed introduction into the linguistics of

grammatical gender and the grammatical gender systems of the languages under investigation: French, Dutch and English. Chapter 5 discusses grammatical gender in CS research and presents the two surveys conducted. These surveys were designed to investigate default gender in (Belgian) Dutch and gender agreement strategies in English-French and Dutch-English CS. Chapter 6 provides a little break, as by that time we will need it.

Grammatical gender is complicated*!*

Part iii focuses on word order. In chapter 7, I discuss verb-second word order in Dutch-English CS. As we will see, the data from CS provide an interesting challenge for the traditional Generative account for verb second. In chapter 8, I investigate placement of the adverb in English-French and Dutch-English CS.

In the conclusion, the results from the different studies that were conducted are summarised. I also discuss the limitations of the dissertation and outline some suggestions for further research.

Part I

BACKGROUND

This part contains all the background information to make sure we are on the same page. Feel free to skip this part if you can't wait to get to the good stuff ☺.

1 THEORETICAL FRAMEWORK

The data collected for this dissertation will be analysed in accordance with Chomsky's Minimalist Program for Generative syntax and Distributed Morphology. This chapter provides a very brief introduction into Generative syntax, the Minimalist Program (MP) and Distributed Morphology (DM). I limit myself here to the fundamental notions, especially those required to understand the analyses that will follow in this dissertation.

For a detailed introduction to Generative syntax, I refer the reader to Den Dikken (2013b). For accessible summaries of the different stages of Generative grammar (and competing frameworks), I refer the reader to Luraghi and Parodi (2008). For a brief, but technical, presentation of DM I recommend Harley and Noyer (1999). A more accessible introduction is provided by Embick (2015).

1.1 THE ORIGINS OF GENERATIVE SYNTAX

Generative syntax has changed a lot since its beginnings. Therefore, before I go on to expand on its most recent iteration, the Minimalist Program, I will provide a brief overview of the origins of the field and the major changes Generative theory underwent.

With the publication of *Syntactic Structures*, Chomsky (1957/2002) founded Generative syntax, though its roots go back to the (American) structuralist movement, which introduced rigorous scientific standards to linguistics. While the structuralist movement itself largely neglected syntax in favour of morphophonemics, Zellig Harris – Chomsky's PhD advisor – argued that any struc-

turalist approach to syntax needed to feature three elements (1), all of which featured heavily in Chomsky's work from the start (Den Dikken 2013a, p 10):

(1) i "statements which enable anyone to synthesise or predict utterances in the language", statements which "form a deductive system with axiomatically defined initial elements and with theorems concerning the relations among them"
ii "statements" which "transform certain sentences of the text into grammatically equivalent sentences"
iii the idea that sentences "consist of a sequence of one or more underlying sentences"

So while Chomsky's Generative grammar is a continuation of the work done in linguistics in the early '50s, it also contrasts with those structuralist approaches. Firstly, Chomsky focused on those aspects of the syntactic system that make it recursive, which structuralism relegated to the realm of performance (Lasnik and Lohndal 2013, p 27). Secondly, Chomsky argued that the main thing to explain about language is what he called "Plato's problem", also known as the "Poverty-of-the-Stimulus argument". Plato's problem is the fact that all children – barring developmental issues – acquire their native language effortlessly, even though what they hear in their environment is "both quantitatively and qualitatively a very poor reflection of what their language is capable of". Structuralist lignuistics, on the other hand, was "very much developed from the linguist's perspective, not the language learner's" (Den Dikken 2013a, p 10–11).

In "Syntactic Structures", Chomsky developed a transformational grammar. This grammar provided a set of phrase-structure and transformational rules that were designed to accomplish the goals stated in (1i) and (1ii). Some examples of phrase-structure rules are shown in (2) (Chomsky 1957/2002, p 26).

(2) a. Sentence → NP + VP
b. NP → D + N
c. VP → V + NP
d. D → THE
e. N → MAN, BALL, ...
f. V → HIT, TOOK ...

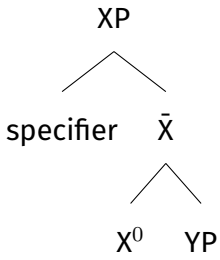
These rules state that the constituent to the left of the arrow can be re-written as the (combination of) constituent(s) on the right side of the arrow. The process of creating a sentence on the basis of these rules is called a derivation. The derivation of THE MAN TOOK THE BALL is shown in (3).

(3) Sentence → NP + VP → D + N + V + NP →
D + N + V + D + N → THE MAN TOOK THE BALL

Transformational rules are rules that apply (optionally) and transform one structure into another. Further development of this idea led Chomsky to propose two levels of representation: Deep Structure (D-Structure) and Surface Structure (S-structure) (Chomsky 1965). While semantic interpretation takes was assumed to take place at the level of D-structure, phonetic interpretation was assumed to take place at the S-structure level. By the publication of *Aspects of the Theory of Syntax* (Chomsky 1965) the importance of the phrase-structure rules was reduced. In the following years, the desire to account for the asymmetrical and hierarchical nature of language led to the development of X-bar theory (Chomsky 1970, 1981).

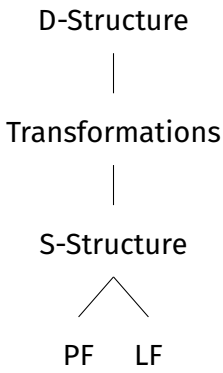
With X-bar theory, Chomsky proposed that every projection is headed by exactly one head. This property is known as endocentricity. Every phrase has the following structure, the $\bar{X}$-schema (4). X^0 is the head of the XPhrase (or XProjection) (Haegeman 1994, p 105).

(4)



X-bar theory figured prominently in Government and Binding (G&B) theory (Chomsky 1981). The foundation of G&B theory is given in (5). This model is known as the (inverted) Y-model or T-model of grammar. While in earlier iterations of the theory, semantic interpretation took place at D-structure, and phonetic interpretation at S-Structure, now both are assumed to take place after S-Structure, in components named Logical Form (LF) and Phonological Form (PF), respectively.

(5)



In G&B theory, syntactic structure is built from the top-down: phrases are projected as ready-made units in the structure before movement takes place. When we turn to code-switching (CS) data, this assumption creates some issues, as I will discuss in section 2.3.

1.2 THE MINIMALIST PROGRAM

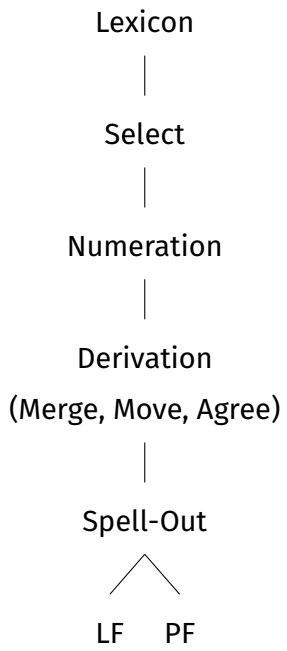
While the MP differs from G&B theory in many respects, a great many of the insights of G&B theory were taken on board in the MP. Three fundamental differences between these approaches are the novel assumption that linguistic variation is situated in the lexicon (6), the elimination of D- and S-structure and the directionality of the derivation.

(6) THE BORER-CHOMSKY CONJECTURE
All parameters of variation are attributable to differences in the features of particular items (e.g. the functional heads) in the lexicon.
Baker (2008, p 353)

Chomsky argues that human language is derived using a single computational system (C_{HL}) and that the variation between individual languages is captured in the lexicon, including whether an element of the synctactic derivation is overtly realised or not (Chomsky 1995/2015, p 7).

At the start of the syntactic derivation, the operation Select takes lexical items from the lexicon and puts them into the Numeration (also known as the Resource or Lexical Array). C_{HL} uses the elements in the Numeration to build sentences, using operations like Merge, Move and Agree. After the derivation is completed, it undergoes the operation Spell-Out and is interpreted at PF and LF (Chomsky 1995/2015, section 4.2.1). This is schematised in (7).

(7)



A well-formed derivation is said to converge, while an ill-formed derivation is said to crash. A crucial notion in Minimalist syntax, inherited from G&B theory, is the functional head. Functional heads are endowed with syntactic features which may trigger movement and "other major syntactic actions" (Rizzi and Cinque 2016, p 141).

Features driving syntactic actions are said to be uninterpretable. They need to be checked against interpretable counterparts in order to be deleted. Uninterpretetable features that are not checked by the end of the derivation cause a derivation crash. Uninterpretable features may be strong or weak. This distinction results in the difference between overt and covert movement. If features are strong, they need to be checked overtly, if they are weak, they may be checked covertly.

As was mentioned above, a final substantial difference between the MP and G&B theory is that the directionality of the derivation is reversed. In the MP the derivation proceeds bottom-up, while in G&B theory, it was a top-down process where phrases

were projected as ready-made units in the structure before movement took place (Luraghi and Parodi 2008, p 31). As we'll see in section 2.3, this shift in directionality is an advantage when it comes to accounting for data from CS.

Another assumption is that clauses have the general structure in (8). The projections in (8) are the complementiser phrase (CP), tense phrase (TP), verb phrase (VP). In earlier versions of the theory, the TP was sometimes referred to as the inflectional phrase (IP). The *v*P is the light verb phrase. As we will see in section 1.3, this *v* is not to be confused with the verbal categoriser in the DM sense. For disambiguation purposes, the light verb phrase is sometimes referred to as *v**P.

(8) $[_{\mathrm{CP}}$ C $[_{\mathrm{TP}}$ T $[_{v\mathrm{P}}$ *v* $[_{\mathrm{VP}}$ V $[\quad]$]]]]

A final component of the MP that needs to be introduced is phase theory. Phase theory is a recent development in the MP, introduced in Chomsky (2000b) and further refined in Chomsky (2008). A phase is defined as a part or chunk of the derivation that has access to only a subset of the Numeration. Crucially, when the derivation of a phase is complete, it is moved to Spell-Out and the phase is no longer accessible to the rest of the derivation. This is known as the Phase-Impenetrability Condition. CP and *v**P are phases. Chomsky (2008) suggests that DPs could also be phases. Crucially, TP is not a phase, an assumption which will be relevant to the operation of Feature Inheritance, which will be introduced in section 7.7.4.

These few elements of phase theory are all we need to know (for now). For a full introduction to phase theory, I refer the reader to Citko (2014). López et al. (2017) discuss the implications of phase theory for CS research, which will be addressed in section 2.5.

1.3 DISTRIBUTED MORPHOLOGY

Distributed Morphology (Halle and Marantz 1993) is an extension of the MP. It is part of a family of approaches whose goal it is to eliminate the distinction between syntax and morphology. The idea that the computational processes that underlie syntax and morphology are one and the same is shared with some other approaches, most notably Nanosyntax (Caha 2009; Starke 2009) and the eXoSkeletal Model (XSM) (Borer 2003). For some key differences between DM and these frameworks, I refer to Borer (2013, section 7.2.2) for XSM and Caha (2018) for Nanosyntax. These approaches are known under the umbrella terms of exoskeletal approaches, late-insertion approaches, or realisational theories of morphology. The following paragraphs are based on Embick (2015), Halle and Marantz (1993, 1994), and Harley and Noyer (1999).

The rationale behind frameworks like DM is the widely held assumption that syntactic terminals are not "words" (Harris 1996, p 99). Even in G&B theory, verbal inflection had already become housed in its own, dedicated functional projection (the tense phrase (TP) or inflection phrase (IP)). DM takes this approach one step further with the idea that syntactic structure extends all the way down to the (lexical) root. Åfarli and Subbarao (2019, p 36) argue that the most important empirical advantage an exoskeletal approach has over lexicalist versions of Minimalism, is that "[exoskeletal approaches] are able to account very neatly for lexical creativity and argument structure flexibility, which is ubiquitous in languages".

Some informally collected examples of such creativity can be found in (9). The context for example (9a) involves characters in a TV show and refers to becoming negatively obsessed with what another person did to you in the past, keeping them in a metaphorical pit in your (mental) basement, as happens unmetaphorically in the 1991 movie *Silence of the Lambs*. Exam-

ple (9b) makes reference to the 2019 Netflix sensation "Tidying up with Marie Kondo". Examples (9c–d) are easily interpretable without any context and were collected in my surroundings.

(9) a. Thomas et al. (2012):
I want to Silence of the Lambs him.

b. Roy (2019):
I Marie Kondo-ed a cluttered classroom, and here's what happened.

c. I've only really twat-ed about today.

d. Ge weet da hij de eikel is omda hij al
you know that he the asshole is because he already
heel de aflevering loopt te eikel-en.
whole the episode runs to asshole-INF
'You know he's the asshole because he's been acting like one for the entire episode.'

DM is named the way it is because in DM the traditional modular conception of morphology has been set aside in favour of a view in which syntactic hierarchical structure goes all the way down and the morphology is *distributed* over several components of the grammar. These components are known in the relevant literature as Lists 1, 2, 3; Lists A, B, C; or the formative list, vocabulary and Encyclopedia (10).

(10) List A: the formative list – stores roots and all feature(-bundle)s that can enter in the derivation
List B: vocabulary – links these roots and abstract features with phonological form
List C: the Encyclopedia – links these features with meaning, and also contains information about idioms

How these three lists feed into the grammar is schematically represented in figure 1.1. The elements from the formative list are directly manipulated by the syntax. They are, in other words, syntactic terminals. DM distinguishes two types of syn-

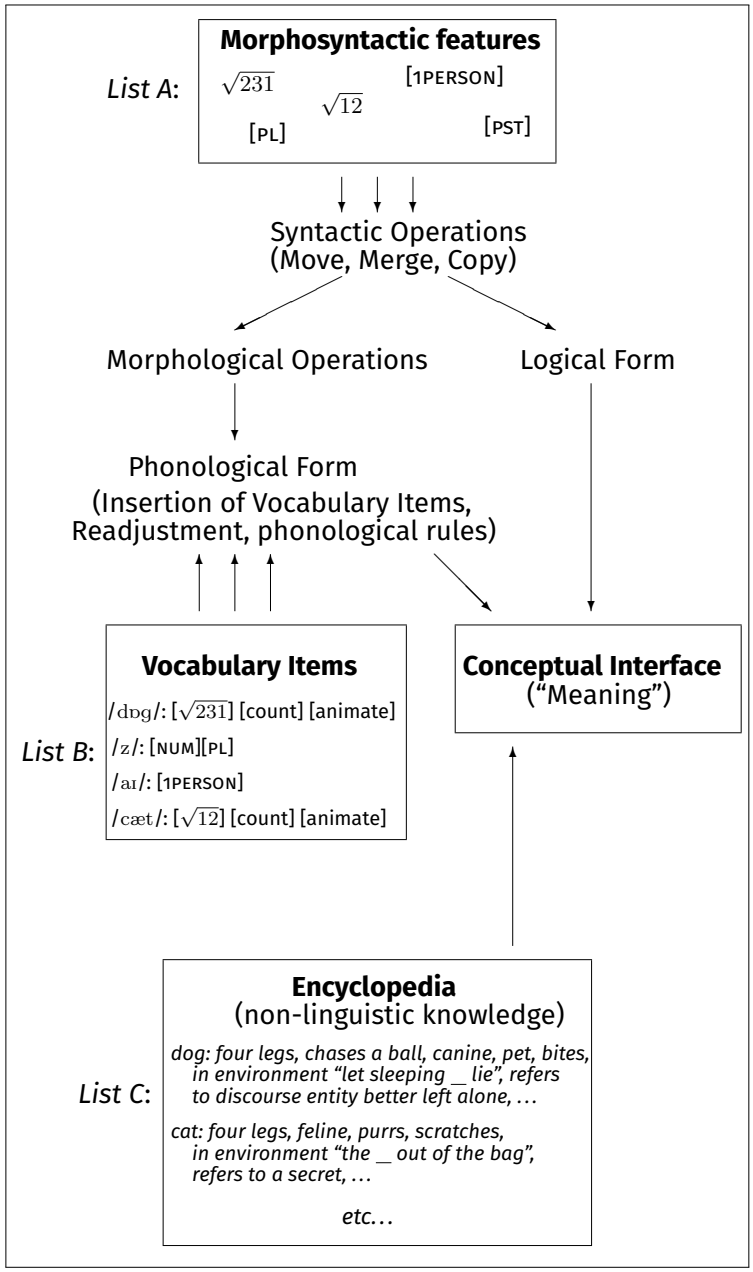
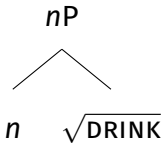


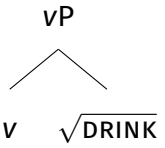
Figure 1.1: Schematic representation of the DM derivation (adapted from Harley and Noyer 1999, p 3)

tactic terminals: roots and functional morphemes. Functional morphemes are composed of syntactic-semantic features, while roots ($\sqrt{\text{ROOT}}$) are devoid of them. Since roots are devoid of grammatical features, it follows that they are by definition categoryless. A root must be merged with a categorising head (such as *n* and *v*), which will result in a noun (11) and verb (12) respectively.

(11) derivation of a noun:



(12) derivation of a verb:



The idea that words consist of a root and a categorising head is also known as lexical decomposition and is a major tenet of modern morphosyntax. The exact nature of roots is a topic of debate in the DM literature (Harley 2014), with some discussion on what information exactly a root can contain. De Belder and Van Craenenbroek (2015) argue that roots are characterised by their position as the element to be subject to what they call "Primary Merge": the first application of Merge in each "cyclic domain" (i. e. basically a word). Similarly, there is some debate about the categorising heads. One of the main differences between DM and XSM is that, in the latter, categorisation of a root is accomplished through Merger with an overt functional head that determines category (such as a number head, or verbal inflection head) rather than through Merger with a categorising head without phonological content (Borer 2014, p 125).[1] While, in classic DM, derivational suffixes are taken to be the exponents of the categorising heads, under some approaches, derivational affixes can be roots themselves (De Belder 2011; Lowenstamm 2012).

1 De Belder and Van Craenenbroek (2015) also take this approach.

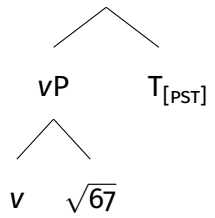
As can be learnt from combining (10) and figure 1.1, a key property of DM is late insertion (13) of vocabulary items.

(13) LATE INSERTION:
The terminal nodes that are organised into the familiar hierarchical structures by the principles and operations of the syntax proper are complexes of semantic and syntactic features but systematically lack all phonological features. The phonological features are supplied – after the syntax – by the insertion of Vocabulary Items into the terminal nodes. Vocabulary Insertion (VI) adds phonological features to the terminal nodes, but it does not add to the semantic/syntactic features making up the terminal nodes.

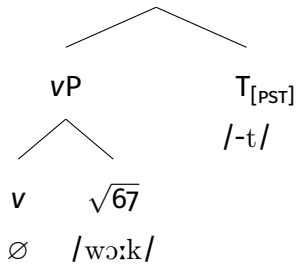
Halle and Marantz (1994, p 275)

This contrasts with early insertion approaches in which word formation is completed before the syntactic derivation begins. In DM, the insertion of vocabulary items only takes place after the derivation is completed. Vocabulary items are a pairing between a phonological exponent and a set of syntactic/semantic features. This means that phonology is entirely absent during the syntactic derivation. (14) and (15) illustrate the derivation of the simple past tense of WALK.

(14) WALKED before VI:



(15) WALKED after VI:



During VI several items may compete for insertion. Which item wins this competition is determined by the Subset Principle (16).

(16) THE SUBSET PRINCIPLE/UNDERSPECIFICATION:
Phonological expressions need not be fully specified for the syntactic positions where they can be inserted. Hence there is no need for the phonological pieces of a word to supply the morphosyntactic features of that word; rather, Vocabulary Items are in many instances default signals, inserted where no more specific form is available.
Harley and Noyer (1999, p 3)

Each terminal node contains a bundle of features and all vocabulary items compete for insertion into the terminal nodes. The vocabulary item that wins the competition is the item that, while containing a subset of the feature specification of that terminal node, is most highly specified for that node. Imagine a syntactic terminal T, specified for features [a,b,c] and some possible candidates for insertion: X[a] and Y[a,b] and Z[a,b,c,d]. In this case, Y wins the competition for insertion, because it is the most specified candidate that is not *over*specified.

These three assumptions – syntactic structure all the way down, late insertion and underspecification – together form the foundation of DM. There is one final distinction that I want to highlight, as it will be relevant in the discussion of grammatical gender in chapter 4: the difference between word derivation and root derivation. Root derived words are words that are derived by the merger of a root with a categorising head (as in (11) and (12)), whereas word derived words are words derived from an already categorised root. One way of distinguishing these two processes is compositionality. Word derivation involves semantic compositionality, whereas root derivation does not. Take the English verbs HAMMER and TAPE. The action expressed by the latter necessarily involves the use of tape (17a), whereas the action expressed by the former does not necessarily involve

a hammer (17b). This indicates that the verb HAMMER is root derived, while TAPE is word derived.

(17) a. *Mac taped their portrait to the wall with a thumbtack.
b. Valda hammered on the door with their fists.

For examples of analyses involving this distinction, see Don (2005) for zero derivation and Lowenstamm (2012) for suffixal derivation.

1.4 SUMMARY

In this chapter, I have very briefly laid out the origins of Generative syntax in general and the Minimalist Program and Distributed Morphology in particular. Section 2.3 in the next chapter lays out the motivations for choosing the MP and DM as the theoretical frameworks in which to analyse CS data. DM will play a role in part ii in particular, as this part contains chapters that concern word formation. As part iii contains chapters that deal with sentence formation, the data will be analysed in a traditional MP framework.

2 LITERATURE REVIEW

Code-switching is the use of more than one language in the same utterance or conversation. Other labels used in the literature for CS and related phenomena are "code-mixing", "code-blending", "language mixing", "language switching", "translanguaging", "codeshifting", "code-changing", "language alternation", "language mixture", "code-meshing", ...

Not all of these labels are used in the same way by all authors. In general, code-switching is the term which is used most often as an umbrella term, even though it is somewhat misleading, as is noted by Gardner-Chloros (2009, pp 11–12):

> Unfortunately, both halves of the term CS *are* misleading. "Code" was originally taken from the field of communication technology [...]. What was meant there has nothing to do with language [...].
>
> "Switching" appears transparent enough, in that it refers to *alternation* between the different varieties which people speak [...] something similar to flicking an electric switch [...]. But accumulated evidence from the mixed speech of bilinguals has led to the *transition* between the two varieties being seen as more and more complex and less and less clear-cut.

The term has, however, been entrenched in the literature and I will be using it as described at the beginning of this chapter. Seminal work by Poplack (1980) on ENGLISH–Spanish CS distinguished three types of CS. Tag-switching (1), inter-clausal (or inter-sentential) CS (2) or intra-sentential CS (3).

(1) vendia arroz 'N SHIT
he.sold rice ...

Poplack (1980, p 589)

(2) SOMETIMES I'LL START A SENTENCE IN SPANISH y termino
... and finish
en Español
in Spanish

Poplack (1980, p 581)

Writing this dissertation was the first time I realised that the title of Poplack's seminal paper (2) has "Spanish" twice ☺.

(3) HE WAS SITTING DOWN en la cama, mirandonos
... on the bed, watching us
peleando, y REALLY, I DON'T REMEMBER si el nos separo
fight, and ... if he separated us

Poplack (1980, p 589)

Muysken (1997) expanded on Poplack's classification and distinguished three patterns of intra-sentential CS: alternation, insertion and congruent lexicalisation. Muysken considers (2) an example of alternation: he claims that "there is a case of true switch from one language to the other, involving both grammar and lexicon" (p 361).[1] Insertion involves the insertion of elements from one language into a structure of another language. We will return to this type of CS in-depth in section 2.2. Finally, congruent lexicalisation "refers to a situation where the two languages share a grammatical structure which can be filled lexically with elements from either language" (p 372). (3) is an example of congruent lexicalisation.

Some authors use the term "code-switching" in a more specific way. Treffers-Daller (1993) for example reserves "code-switching" for inter-sentential CS, while she calls intra-sentential CS "code-mixing". Meyerhoff (2018, chapter 6) uses a similar distinction, but seems to include "language choice" (i. e. deciding which variety to use for which interaction) under the umbrella of CS. Some authors avoid the term CS altogether, preferring a more

1 He seems to be saying that alternation is a possible pattern for both inter- and intra-sentential CS, as he says it takes place "between utterances in a turn or between turns" (p 361).

transparent term like "language mixing" (e.g. Åfarli and Subbarao 2019; Grimstad et al. 2014).

"Code-blending" is usually used for a subtype of CS and denotes the mixing of a signed and a spoken language. Since the two languages involved in code-blending make use of different modalities, it does not necessarily involve alternation between the two languages, but parts of the utterance may be produced in both modalities at the same time, or partially in one and partially in the other with varying degrees of overlap. For an introduction to code-blending see Hill et al. (2018, section 5.2).

The final term I want to mention is translanguaging. This term is a translation of the Welsh "trawsiethu" coined by Williams (1996). Originally Williams used it to refer to a pedagogical practice in which students alternate languages when moving from the perceptive to productive mode (Pennycook 2016, p 203). They might, for example, read a novel in English and provide a summary in Welsh. However, the term soon transcended this somewhat narrow definition and has now "emerged as the term of choice [...] referring to instances of language mixing" (Pennycook 2016, p 202). I have the impression that, although all instances of CS fit within this broadened conception of translanguaging, not all translanguaging would be considered CS. The translanguaging movement is also "a sociopolitical attempt to valorise the practices of multilingual speakers" (Meyerhoff 2018, p 136), as within the movement the term "CS" is seen as having acquired negative connotations. For a thorough overview of terminology in this translanguaging movement, I refer the reader to Pennycook (2016).

2.1 EARLY APPROACHES

The study of CS took off as a separate field in the early 1970s (Appel and Muysken 1987, p 117). However, Benson (2001) presents

a couple of studies from much earlier that have been largely forgotten in the literature. Though the term "code-switching" has been in use in the literature only since 1954 (Benson 2001, p 26), the earliest study in the phenomenon she found was a 1909 PhD dissertation investigating the Hispanic communities in New Mexico and Colorado. This dissertation was published in several parts, and the third part, entitled "The English Elements", was published in 1914 (Espinosa 1914). Espinosa published on the subject between 1911 and 1917 noting CS as a pervasive phenomenon in the communities he studied, intersecting all levels of education and social class. The next code-switching study dates from nearly 40 years later (Barker 1947) and investigated the linguistic and social behaviour of Mexican Americans in Tucson, Arizona (Benson 2001, p 28). In this work, Barker notes motivations for CS, such as topic of conversation, group membership and interpersonal relations.

to me at least!

That these early investigations focused on the motivation for CS is unsurprising as the question of why bilinguals mix their languages is probably the number one question I have been asked when I am discussing my research with people who aren't well-versed in the topic. Between 1914 and now, a large amount of factors that motivate CS have been proposed in the literature. Gardner-Chloros (2009, section 3.2) divides those factors into three groups, though these groups only provide " a semblance of order within the huge range of factors" and there is some degree of overlap between them (p 43).

- factors independent of particular speakers/circumstances, which affect all the speakers of the relevant varieties in a particular community, for example:
  - economic factors
  - (c)overt prestige
  - power relations
  - associations of each variety with its own context
- factors dependent on the speakers, for example:

  - competence in the varieties involved in CS
  - social networks of the speakers
  - attitudes and ideologies of the speakers
- factors within the conversations, for example:
  - quoting
  - topic of of the conversation

Appel and Muysken (1987) remark that the functions of CS do not differ qualitatively from the functions of "monolingual" language use.[2] Indeed, all of the motivations listed above sound very familiar to anyone that has read an introductory textbook on sociolinguistics (see for example Meyerhoff 2018; Trudgill 2000). Though they are interesting and well worth investigating, it is not within the scope of this dissertation to provide a detailed overview of the functions proposed in the literature. I refer the reader to Auer (1988) and Blom and Gumperz (1972) for seminal work on the sociolinguistic functions of CS, Appel and Muysken (1987, section 10.1) and Gardner-Chloros (2009, chapter 3) for an overview of the functions proposed in the literature at that point, and James (2016) and Lee (2015) for a more recent approach to the functions of translanguaging and CS extending beyond traditional discourse/conversation analysis.

Returning to the early study of CS, the first works dealing with the phenomenon that are often cited in the field today are Haugen (1953a,b) and Weinreich (1953) and the term CS was first recorded in a review (Vogt 1954) of the latter (Benson 2001, p 26). Though these works are seminal in the field of language contact, there is comparatively little space devoted to CS, and the little there is is extremely negative. Haugen attributed CS to either insufficient knowledge of a second language or "individuals of low-grade intelligence" (1956, as cited in Benson 2001, p 24), while

2 Based on the work of Jakobson, they propose six functions of the language system, all of which apply to CS as well: the referential, directive, expressive, phatic, metalinguistic and poetic functions (Appel and Muysken 1987, section 10.1).

Weinreich noted that "the ideal bilingual switches [...] certainly not within a single sentence" (1953, p 79).

While Espinosa did not consider CS to be a deviant phenomenon, he did observe an apparent lack of grammatical rules:

> The kind of speech mixture which brings into the Spanish of New Mexico the use of regular English words and phrases has no fixed limits and cannot follow regular law.
>
> Espinosa (1914), as cited in Benson (2001, p 31)

The idea that CS is an unregulated language mixture motivated by lack of competence in (one of) the languages involved was probably the commonly held view until at least the mid-seventies, as is reflected in this 1971 quote by Labov about Spanish-English CS:

> So far, however, no one has been able to show that such rapid alternation is governed by any systematic rules or constraints, and we therefore must describe it as the irregular mixture of two distinct systems.
>
> Labov (1971, p 457)

Unfortunately I have not been able to unearth a copy...

So many female pioneers!

While the first study of the grammar of CS seems to be Lehtinen (1966), a PhD dissertation investigating an English-Finnish bilingual corpus, the second part of which is entitled "Rules for Code-Switching and Code-Accommodation", the first published articles on the topic all date from the mid-to-late 70s and all focus on Spanish-English CS in different communities (Aguirre 1976; Gingràs 1974; McClure 1977; Pfaff 1979; Timm 1975). Poplack's seminal paper "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching" was probably the paper that brought research on the grammatical constraints of CS into the mainstream (Poplack 1980).

Poplack (1980) is well-known for the two constraints she postulated: the Equivalence Constraint (4) and the Free Morpheme Constraint (5). Versions of this former constraint had already

been posited by Lehtinen (1966, p 153)[3] and Pfaff (1979, p 314), among others.

(4) THE EQUIVALENCE CONSTRAINT:
Surface structures common to both languages are favored for switches

Poplack (1980, p 583)

(5) THE FREE MORPHEME CONSTRAINT:
Word-internal CS is not allowed

Poplack (1980, p 585)

The ban against word-internal switching (5) is heavily debated in the literature in the context of distinguishing CS from borrowing. I will return to this issue in section 2.5. As for the equivalence constraint, research into CS at points where the surface structures of the languages involved are not equivalent, known as conflict sites, has proliferated in the last 15 or so years, in particular in the Generative approach to CS.

A final constraint that was posited in this early literature on the grammar of intra-sentential CS was that pronominal subjects or objects must be realised in the same language as "the verb to which they belong" (Timm 1975, p 477). I believe there is a consensus that this particular constraint holds, though see González-Vilbazo and Koronkiewicz (2016) for some nuance.

In the next section, I discuss an influential approach to CS, the Matrix Language Frame model. In the subsequent sections (sections 2.3 and 2.4) I will return to the Generative approaches to CS.

3 "In order for any intra-sentence code-switching to be possible at all, there must exist in the two languages some constructions which are in some sense similar, so that certain syntactic items from each language are equivalent to each other in specific ways. Further reflection, supported by an examination of the corpus, shows that the similarities must exist in what is known as the 'surface grammar' of sentences".

2.2 THE MATRIX LANGUAGE-FRAME MODEL

Among those models that aim to capture the grammatical constraints on CS, the Matrix Language Frame (MLF) model (Myers-Scotton 1993/1997) is the most influential one outside the Generative approaches. While I will not be using the MLF model in this dissertation, it is so prevalent in CS research that a brief summary of the framework is warranted.

The MLF model is a formalisation of the observation, credited to Joshi (1985), of an asymmetry in CS: in – what Myers-Scotton calls – "classic" CS,[4] the grammatical structure of a clause comes from only one of the participating languages. This language is called the "Matrix Language", and it is in the Matrix Language (ML) that the other language(s), known as the "Embedded Language(s)" (EL), are embedded (Myers-Scotton 2006, p 243). Several important principles are at work in this model (definitions are taken from Myers-Scotton 2006, pp 243–244):

- the Uniform Structure Principle (USP):
  A given constituent type in any language has a uniform abstract structure and the requirements of well-formedness for this type must be observed whenever the constituent appears. In bilingual speech, the structures of the Matrix Language are always preferred, but some Embedded structures [...] are allowed if Matrix Language clause structure is observed.
- the Morpheme Order Principle (MOP):
  In mixed constituents consisting of at least one EL word and any number of ML words, surface word (and morpheme) order will be that of the ML.

4 This contrasts with "composite" CS, which occurs less frequently, and is defined as "bilingual speech in which even though most of the morphosyntactic structure comes from one of the participating languages, the other language contributes some of the abstract structure underlying surface forms in the clause" (Myers-Scotton 2006, p 242).

- the System Morpheme Principle (SMP):
  In [...mixed] constituents, all system morphemes [...] will come from the ML.

While the first one of these also applies to monolingual speech, EL-islands, which are chunks of EL, form an exception to the SMP and MOP, and are allowed if they meet EL well-formedness conditions, as well as those ML conditions applying to the clause as a whole, such as phrase placement (Myers-Scotton and Jake 2009).

In order to account for the morphemes that are targeted by the SMP, Myers-Scotton (2006, p 244) developed the 4M model of morpheme classification. This model (revised in Myers-Scotton and Jake 2017) first makes a distinction between content and system morphemes. System mophemes are divided into early and late system morphemes, the latter of which are again divided into two types: bridge and outsider system morphemes. Only the late outsider system morphemes are targeted by the SMP (pp 342–344).

Content morphemes are the "main conveyors of semantic meaning" (Myers-Scotton and Jake 2017, p 343). Nouns (I think they mean nominal stems) are a clear example of such morphemes. Early system morphemes are grammatical morphemes whose meaning is nonetheless "conceptually salient" (p 344), such as the plural morpheme (on nouns), or derivational affixes. Late system morphemes build hierarchical grammatical structure and don't add any semantic meaning. Bridge system morphemes may add pragmatic meaning. These morphemes, as the name suggests, join together two units. The preposition OF in THE COVER OF THE BOOK or the complementiser THAT in I EXPECT THAT YOU WILL DO IT are examples of bridge system morphemes (p 345).

Finally, outsider morphemes depend on information that is outside the word to which the morpheme attaches. This information can be retrieved from another word in the utterance or from

the discourse. Verbal agreement and case marking are prime examples of late outsider system morphemes (Myers-Scotton and Jake 2017, p 346).

While in early iterations of the MLF, the ML was determined at corpus-level by looking at the language which is the most frequent in a corpus (Myers-Scotton 1993/1997, p 68), in later versions the ML is determined *both* at the level of the corpus and at the level of the clause, by looking at which language determines the word (or morpheme) order and which language provides the system morphemes, in accordance with the SMP and MOP (Myers-Scotton and Jake 2009, p 338). In many recent approaches (e.g. Blokzijl et al. 2017; Herring et al. 2010; Vanden Wyngaerd 2017), the ML of a clause is determined by looking at the language which provides the inflected verb, as verbal agreement is a (perhaps *the*) prime example of a late outsider system morpheme. This despite the fact that Myers-Scotton (1993/1997, p 66) explicitly argues against defining the ML as the language of the inflected verb: "to avoid circularity, the ML must be defined independent of this structural role which it plays".

The MLF model is notably quite successful in predicting "co-occurrence restrictions within the clause" (Deuchar 2020, p 16) such as the mixed-DP asymmetry: in CS corpora, the ML tends to provide the determiner in mixed determiner phrases (DPs). This asymmetry is not necessarily easy to account for within a Generative approach to CS (Jake et al. 2002, p 70). This issue will be discussed in detail in section 5.1.1.

2.3 GENERATIVE APPROACHES TO CS

While the early approaches discussed in section 1.1 could perhaps fall under the Generative umbrella, the first explicit Chomskyan approaches to CS date from the G&B era. Well-known studies within G&B theory include Belazi et al. (1994), Bentahila and Davies

(1983), Di Sciullo et al. (1986), Joshi (1982), and Woolford (1983). With this new wave in Generative approaches to CS, the focus shifted from postulating constraints that focus on one structure or constituent to postulating formalised constraints that were applicable more generally (Di Sciullo et al. 1986, p 3). While Woolford (1983) tried to develop a mechanism that accounted for the Equivalence Constraint (4) within G&B theory, most of these G&B approaches moved away from the Equivalence constraint. Belazi et al. (1994), Bentahila and Davies (1983), and Di Sciullo et al. (1986) note that it is empirically inadequate, with the latter adding:

> A second general problem with the equivalence constraint is that it is formulated exclusively in terms of linear sequence, rather than in terms of structural relations. Since we hold that most principles of grammar are formulated in terms of hierarchical relations rather than of linear order, and since code-mixing appears to involve central aspects of grammatical competence, it would be necessary from the point of view of the theory of grammar that constraints on code-mixing are structural rather than linear.
>
> Di Sciullo et al. (1986, p 3)

Some constraints that were posited in approaches of that kind are listed in (6), (7) and (8).

(6) THE GOVERNMENT CONSTRAINT:
if X has language index *q* and if it governs[5] Y, Y must have language index *q* also

Di Sciullo et al. (1986, p 228)

5 The definition Di Sciullo et al. (1986) use is: X governs Y if the first node dominating X also dominates Y, where X is a major category (N, V, A, P) and no maximal boundary intervenes between X and Y.

(7) THE FUNCTIONAL-HEAD CONSTRAINT:
The language feature of the complement f-selected by a functional head, like all other relevant features, must match the corresponding feature of that functional head.
Belazi et al. 1994, p 228

(8) CONSTRAINT ON CLOSED CLASS ITEMS:
Closed-class items (e.g. determiners, quantifiers, prepositions, ...) cannot be switched.
Joshi 1982, p 148

From the earliest stages in Generative approaches to the grammar of CS, there has been a desire to avoid CS-specific mechanisms. Pfaff (1979, p 314) for example argued that "[i]t is unnecessary to posit the existence of a third grammar to account for the utterances in which the languages are mixed". Di Sciullo et al. (1986, p 7) note: "code-mixing can be seen as a rather ordinary case of language use, requiring no specific stipulation". Joshi (1982, p 150) claims that "there is no third grammar" that regulates CS.

Approaches that do not postulate CS-specific mechanisms are also known as constraint-free approaches. This label does not indicate that such approaches assume that CS is unconstrained, but rather that they are constrained by nothing except the grammar of the languages involved in CS. However, the tools to develop an approach that does not rely on CS-specific mechanisms were not really available before the advent of the MP. The constraints listed in (6)–(8), all take recourse to mechanisms that rely on the notions of "language label" or "language switching" explicitly.

Mahootian (1993)[6] is often cited as the first one to have developed a true constraint-free approach to CS (which she called a "null theory"). Her approach was developed within the Tree-Adjoining Grammar (TAG) formalism (Joshi 1987), a generative formalism of syntax roughly consistent with the principles of G&B

6 I would like to thank Dr Mahootian for sending me a copy of her dissertation.

theory (Mahootian 1993, p 141). A crucial difference between TAG and more mainstream G&B approaches is the increased importance accorded to the lexicon. This importance of the lexicon is a property TAG and the MP have in common. As we have seen in section 1.2, in the MP language consists of two components: the computational system and the lexicon. The former is universal and invariant, while the latter is language-specific and contains the parameters responsible for the wide variation attested within the world's languages (MacSwan 2009, p 321).

Reducing linguistic variation to the lexicon – i.e. the lexical items with their language-specific syntactic features – allows for a view of bilingualism in which the (morpho-syntactic) grammars of the two languages are less compartmentalised than in earlier iterations of Generative theory, such as G&B theory (MacSwan 2009). This has made accounting for CS phenomena easier. As was explained in section 1.1, before the MP, lexical insertion was postponed "until well after the word order had been laid out, posing a significant problem. The structure could not be sensitive to which language contributed a specific lexical item until the end, when lexical insertion occurred, but the language contributing the lexical item appeared to have strong consequences for the syntactic structure at the onset" (MacSwan 2009, p 320).

This significant problem has disappeared, as structures are now considered to be built from lexical items *up*. This accounts for the influence that lexical items – and the language they belong to – have on the grammatical structures. The first one to develop an account of CS within the MP was MacSwan (1999). In figure 2.1 his model for CS is represented. In code-switched sentences, items from Lexicons α and β are put in the Numeration via the operation Select. As we saw above, the computational system (C_{HL}) then uses the elements in the Numeration to build sentences.

MacSwan's work was quite influential and research into what happens when CS occurs at so-called "conflict sites" increased.

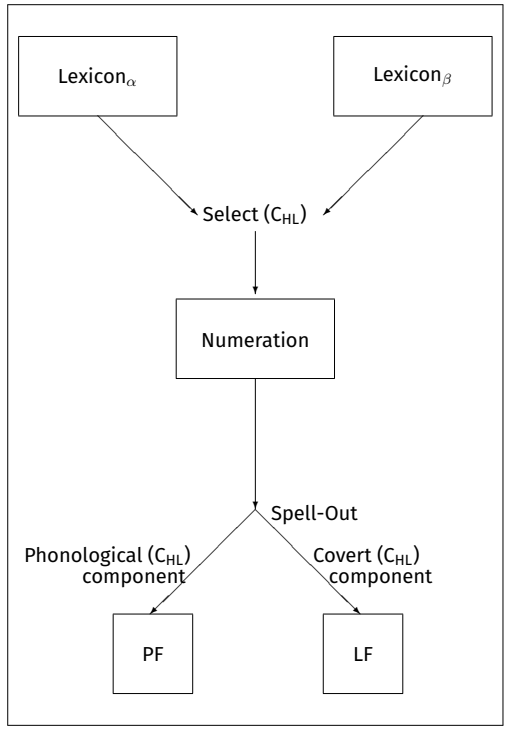


Figure 2.1: The Minimalist Framework: derivation of CS sentences. Adapted from MacSwan (2009, p 179)

Conflict sites are structures for which the properties of the two participating languages differ. Word order is a popular example of a conflict site (Cantone and MacSwan 2009; González-Vilbazo and López 2012; Herring et al. 2010; Pablos et al. 2019; Shim 2013; Stadthagen-González et al. 2018; Vanden Wyngaerd 2017). Chapters 5, 7 and 8 include a detailed review of the CS-literature on the relevant conflict sites.

2.4 DM APPROACHES TO CS

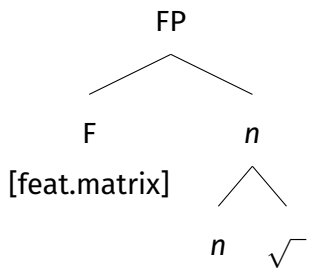
While DM (Halle and Marantz 1993) is about the same age as the MP, research on CS invoking DM mechanisms is a relatively recent phenomenon. The first published paper – to my knowledge

– that explicitly acknowledges extending MacSwan's proposal and incorporating a DM approach to the lexicon is Liceras et al. (2008), though the analysis presented there does not go below the word level. More recently there have been a spate of papers that specifically advocate abandoning lexicalist models like MacSwan's in favour of an exoskeletal approach like DM (see for example Grimstad et al. 2014; López 2020; Riksem et al. 2019).

Under such approaches, the languages involved in CS are even less compartmentalised than in a model like MacSwan's. Recall from section 1.3 that DM has no lexicon in the traditional sense. The information that is usually considered to be in the lexicon is spread over three lists. While under MacSwan's lexicalist approach there are two lexicons that each store the relevant lexical items, in DM approaches to CS, items from the languages involved in CS are stored alongside each other in these respective lists.

Åfarli and Subbarao (2019)[7] develop a specific model to account for CS in a DM framework: the Exoskeletal Frame Model (EFM). In the EFM, words (in the traditional sense) consist of fixed morphosyntactic frames. Such frames consist of a "backbone" of abstract functional heads, which are composed of feature matrices. The functional heads in the morphosyntactic frames tend to be provided by one language, though specifiers and adjunct positions are much more "liable to insertion [...] of material from other languages" In (9) the morphosyntactic frame for a noun is given, with F standing for an abstract functional head (Åfarli and Subbarao 2019, p 37–38).

(9)



7 I would like to thank Dr Åfarli who sent me a copy of this chapter.

A concrete example of a noun in French (10a), a language with gender and number, together with its frame (10b) is given below.

(10) a. cousin-e-s
cousin-F-PL
'female cousins'
b. [$_{\text{NumP}}$ [$_{\text{GenP}}$ [$_{n\text{P}}$ *n* [$\sqrt{\text{COUSIN}}$]] -e] -s]

The EFM is an attempt to incorporate some of the success of the MLF model into Generative approaches. Recall that the MLF is notably more successful than Generative approaches in capturing co-occurrence restrictions. While the MLF and Generative approaches have very different theoretical underpinnings, Åfarli and Subbarao (2019) argue that the EFM and MLF have the following two elements in common (11):

(11) i that frames, corresponding to the MLF's ML frames, are generated independently of lexical items, and
ii that lexical insertion, e.g. insertion of exponents, takes place late in the derivation (Late Lexical Insertion)
Åfarli and Subbarao (2019, p 37)

A similar approach is developed by López (2020) and López et al. (2017), the latter of which also incorporates phase theory. López et al. (2017) will be discussed in more detail in section 2.5. Section 5.1.1 discusses DM approaches to CS in more detail, in particular analyses for grammatical gender that have been developed in the literature.

While the MP and DM are compatible with each other, in the recent CS literature a distinction is often made between "lexicalist" Minimalist approaches to CS (such as MacSwan's approach) and DM approaches. In lexicalist approaches, a strict separation between the two lexicons is maintained.

2.5 CS VERSUS BORROWING

Though many credit Weinreich (1953, section 2.4) with first drawing a distinction between borrowing, which he described as "outright transfer of the phonemic sequence from one language to another" (p 47), on the one hand and CS on the other, Benson (2001) notes that Espinosa (1914) already made that distinction.

How exactly to classify lone other-language items (LOLIs) is the topic of a good amount of debate in the CS literature. This question is not only of a theoretical nature, but also has some practical consequences. The language membership of lexical items determines their feature inventory, and consequently their behaviour, so determining the linguistic identity of these words is a relevant question.

In her review of the literature, Grimstad (2017, p 3) identifies two broad positions on the classification of LOLIs (12).

(12) a. Borrowing is the diachronic process by which languages enhance their vocabulary (or other domains of structure), while code-switching consists in instances of spontaneous language mixing in the conversation of bilinguals. Borrowed items originate as code-switches.

b. Code-switching involves inserting alien words or constituents into a clause; borrowing involves entering alien elements into a lexicon.

Poplack (2018) is an example of the position summarised in (12b). For Poplack, the hallmark of borrowing is the morphosyntactic integration of a LOLI. If this LOLI is established in a community, it is considered a loanword. If the LOLI is not established in the community, it is considered a "nonce borrowing" (Poplack et al. 1988). It is borrowed, if only for the nonce. This category of "nonce borrowing" allows Poplack to account for apparent counter-examples to the Free Morpheme Constraint. Synchron-

ically, she argues extensively, these nonce borrowings behave exactly like established loanwords. She argues that "loanwords do not originate as code-switches" but rather through nonce borrowing (Poplack 2018, p 213). LOLIs that are true code-switches (i. e. the insertion of alien words into a clause, cf (12b)) are quite rare under Poplack's approach.

In the discussion of borrowings versus code switches an important issue is the ban on word-internal code-switches, i. e. Poplack's Free Morpheme Constraint (5). Many approaches to CS adopt the Free Morpheme Constraint (MacSwan 1999; MacSwan and Colina 2014 among others). Others, however, argue that CS within the word is possible and does occur (Bolonyai 2005; Deuchar 2020; Jake et al. 2005 among others).

In Minimalist approaches to CS, the Free Morpheme Constraint is currently formalised as the PF Interface Condition (PFIC) (MacSwan and Colina 2014) – formerly known as the PF Disjunction Theorem (MacSwan 1999). The PFIC encapsulates the idea that the nature of phonological processing is such that switching from one phonological system to another word-internally is not possible (13).

(13) PF Interface Condition (MacSwan and Colina 2014, p 191)

i Phonological input is mapped to the output in one step with no intermediate representations.

ii Each set of internally ranked constraints is a constraint dominance hierarchy, and a language-particular phonology is a set of constraint dominance hierarchies.

iii Bilinguals have a separately encapsulated phonological system for each language in their repertoire in order to avoid *ranking paradoxes*, which result from the availability of distinct constraint dominance hierarchies with conflicting priorities.

iv Every syntactic head must be phonologically parsed at Spell-Out.

Readers familiar with Optimality Theory (OT) (Prince and Smolensky 1993) will recognise this conception of phonology. In this Generative model of phonology, the tension between faithfulness to the underlying representation and markedness is central. The phonology of a language is no more (or no less) than the ranking of the set of universal phonological constraints, some of which favour markedness, others faithfulness.

Looks like those phonology courses I took are coming in handy ☺!

The difference between the pronunciation of the final consonants of English BID /bɪd/ and Dutch **bied** /biːt/ 'bid' is due to a difference in ranking between the faithfulness constraint IDENT$_{(VOICE)}$ (preserve voicing specification) and the markedness constraint *VC (no voiced obstruents in the coda) (Itô and Mester 2003). Take a Dutch-English bilingual who tries to switch word-internally from the Dutch past participle prefix **ge**- to the English stem BID. There will be a conflict in constraint-rankings for the prefix and stem. Since **ge**BID needs to be phonologically parsed at Spell-Out (13iv) the conflict in constraint-ranking means the form cannot be pronounced, according to the PFIC. Note that the PFIC does not rule out a switch like **ge**BID, provided that it is pronounced as either a Dutch or an English word in its entirety, though it makes no predictions – as far as I know – as to which pronunciation would be selected. The empirical validity of the PFIC has also been called into question, with many authors providing counter-evidence (see for example Bandi-Rao and Dikken 2014; Bolonyai 2005; Deuchar 2020; Grimstad 2017; Jake et al. 2002, 2005).

As was discussed in the previous chapter, DM does away with the traditional distinction between syntax and morphology. Words are formed using the same syntactic operations that are used to form clauses. Hence it not surprising that DM approaches to CS do not assume switching within the word to be impossible. In addition, in contrast with lexicalist Minimalist approaches,

under DM, there is no longer a separation between the different lexicons of the multilingual individual. Hence, neither of the two approaches outlined in (12) to distinguish CS from borrowing can be adopted by adherents to DM.

Grimstad (2017) replaces the opposition between CS and borrowing by an opposition between "insertion" and "alternation", based on Muysken's classification of CS. The former involves a LOLI being inserted with recipient language inflection, cf Poplack's (nonce) borrowings, whereas the latter consists of a LOLI being inserted with its own inflection, with alternations being quite rare. In other words, insertion is the inclusion of an "other-language" root in a morphosyntactic frame, whereas alternation is the inclusion of an entire "other-language" morphosyntactic frame, including its root.[8]

However, it is not the case that, because switching within the word is possible, "anything goes". Bandi-Rao and Dikken (2014) use data from Telugu-English CS to argue that whether or not CS within the word is possible depends on the nature of the affix. If the affix incorporates its complement, the result is a morphosyntactic head and CS between the root and affix is illicit. If the affix does not incorporate its complement, the result is not one morphosyntactic head. Rather, it is a complex, polymorphemic unit and CS between head and affix is possible.

López et al. (2017), incorporate insights from phase theory. For example, they observe that switching between a root and affix is allowed, but switching between a derivational and inflectional affix is ruled out. Case morphology provides an intriguing exception to this generalisation. Recall from section 1.2 that in phase theory, each phase is transferred to Spell-Out as soon as the derivation has completed building that phase. The reason why switching between derivational and inflectional suffix is not

8 I have put "other language" between scare quotes, not to suggest that there is any grammatical feature that represents the language membership, but rather that these elements are traditionally considered as belonging to another language.

allowed is because these suffixes are transferred to Spell-Out at the same time. The root however, is transferred independently of the affixes.[9] The approach taken by López et al. also accounts for the phonology of words containing word-internal mixes. Take a look at the example in (14), containing **German**-Spanish word-internal CS.

(14) cabre-**ier**-**t**
anger-V-PTCP
'angered' López et al. (2017, p 11)

In this example, we have a Spanish root $\sqrt{\text{CABRE}}$ ('anger'), which is merged with the German derivational morpheme -**ier**, which derives a verb. Since the categorising head is the phase head, and the phase head determines the grammatical properties of the phase, it is the language of the derivational morpheme that determines the phonology of the word it derives. That means that the word in (14) is pronounced with German phonology (López et al. 2017, p 13–15). The PFIC is unable to make predictions of this kind.

Overall, I think approaches such as the ones developed by Åfarli and Subbarao (2019) and López et al. (2017) are promising. It seems that there are regularities that underlie multilingual word-formation, just like there are regularities that underlie multilingual sentence-formation, which lexicalist Minimalist approaches are unsuited to capture. However, at this point, these DM approaches are somewhat underdeveloped. It is unclear, for example, how the ban on switching between derivational and inflectional affixes is "enforced". This intuition seems – to me at least – legitimate, the underlying theoretical mechanism is clear (transfer in the same phase), but it is not explicit how exactly this switch is prevented.

9 The reason why case morphology is an exception is because the head that houses the case morphology transfers in a different phase than its complement (López et al. 2017, p 13).

2.6 PSYCHOLINGUISTIC AND EXPERIMENTAL APPROACHES TO CS

In the last decade or two, a growing body of literature has addressed questions of bilingual language processing by approaching CS from a psycho- and neurolinguistic perspective. The focus of earlier studies in the field used "language switching" tasks, which were mainly based on the task-switching paradigm that is popular in cognitive psychology. However, such switching tasks are not exactly comparable to the CS bilinguals engage in in natural discourse (Van Hell et al. 2018, p 135).

More recently however, "[a]n emergent body of studies seeks to examine the cognitive and neural correlates of language switching in more naturally occurring situations: switching within meaningful sentences, i.e. intra-sentential code-switching" (Van Hell et al. 2015, p 462). Such studies contribute to our understanding of the multi-faceted phenomenon that is CS.

Unfortunately, an outline of the contributions and insights of this research field is outside the scope of this dissertation, and I will limit myself to making some recommendations for further reading. Gullberg et al. (2009, section 2.3) provides a review of experimental and neurocognitive methods and tasks that are used in the field. A more elaborate overview of neurocognitive methods and their benefits can be found in Van Hell et al. (2018). Some interesting studies on syntactic co-activation include Luque et al. (2018), Sanoudaki and Thierry (2014), and Vaughan-Evans et al. (2020). For examples of how insights from corpus-research can be integrated with those of neurocognitive approaches I recommend Beatty-Martínez et al. (2018) and Valdés Kroff et al. (2018).

2.7 SUMMARY

In this chapter, I have provided a concise account of the general literature on CS, with a particular focus on the frameworks that will be used in this dissertation. This overview provides the backdrop for the topics that will be presented in parts ii and iii of this dissertation. Chapters 5, 7 and 8 include a more detailed review of the literature concerning these topics. In the next chapter, I will discuss some methodological issues, before moving on to parts ii and iii.

3 METHODOLOGICAL CONCERNS

3.1 ACCEPTABILITY JUDGMENTS

> I have yet to see a descriptive grammar of a language that is not entirely built on native speaker intuitions.
>
> — López (2020, p 10)

The intuitions of native speakers have long formed the foundation of linguistic research. These types of intuitions are often referred to as grammaticality or acceptability judgments. As Schütze and Sprouse (2014, pp 27–28) point out, the former label is misleading. Grammaticality is, strictly speaking, a mental construct inaccessible to the awareness of speakers. Asking people to judge whether a sentence is "good" or "bad" is asking them about the acceptability of said sentence.

Acceptability judgments rely on the assumption that acceptability is a "percept that arises (spontaneously) in response to linguistic stimuli" (Schütze and Sprouse 2014, p 28). Just like other experiences like pain and temperature, there is no direct way to measure how participants experience such percepts. A common method in the field of perception research is asking participants to report their perceptions on a scale. This is a method that has been widely applied in linguistics as well. Most syntacticians assume that acceptability judgments can be used as evidence for "the grammatical system of the human language faculty" (Schütze and Sprouse 2014, p 28). For more on the importance of acceptability judgments in the field of Generative syntax, I refer the reader to Sprouse (in press), a chapter specifically geared towards linguists that prefer not to use acceptability judgments.

While acceptability judgments are widely used in linguistic research, they are not uncontroversial. Labov (1972) dedicates a subsection to the "problems in the study of intuitions", in which his main argument is that it is not the case that "the great majority of these [intuitions] would be clear judgments" (p 159). Acceptability judgments have been highly criticised over the last half century or so for a variety of reasons. The main criticism levelled at judgment data revolves around the reliability of such data when they are informally collected, either because such judgments rely on the intuitions of a limited number of people, or on the intuitions of the author(s) of a paper, who may be susceptible to (unconscious) confirmation bias (Gross in press, p 8). Recently, however, there have been a few replication studies that show that such informal judgments have a very high rate of replicability – between 95 and 98% – with large sample sizes of naive native speakers (Sprouse and Almeida 2012; Sprouse et al. 2013).

Acceptability judgments can also be collected more formally, through an acceptability judgment task (AJT). In such a task a larg(er) number of usually naive speakers are asked to provide an acceptability judgment for certain strings of words.[1] A criticism often brought up in the context of such tasks is that we do not know what is being measured in an AJT task. Firstly, there is a concern that the phrasing of the instructions may have unintended influence on how participants perform in an AJT (e.g. Branigan and Pickering 2017, p 4), though Cowart (1997, pp 56–59) showed via an experiment where he manipulated instructions that, as long as they get at least some explicit instructions, participants in different groups did not rate sentences differently. Secondly, some argue that the source of acceptability may be related to factors beyond grammaticality. Indeed, Sprouse (in press, section 2.2) notes:

1 For a detailed breakdown of the construction of such tasks, I refer the reader to Cowart (1997).

> I think all syntacticians assume that there are multiple factors that impact that error signal – grammar (phonology, morphology, syntax, semantics, pragmatics, etc), language processing (parsing strategies, working memory, predictive processes), real world knowledge (e.g. plausibility), task effects, etc.

However, "grammatical effects are likely to have larger effect sizes than extra-grammatical effects" (Sprouse and Almeida 2017, p 27), which leads most Generative syntacticians to assume that the mapping between grammaticality (i. e. syntactic well-formedness) and acceptability judgments is relatively direct.

The controversy that surrounds acceptability judgments in the monolingual literature is compounded in the CS literature. There are two main factors that come into play. Firstly CS is considered to be an oral phenomenon by many (Mahootian 2005; Montes-Alcalá 2000; Poplack 1980), meaning that written AJTs are not suited to investigate CS. However, there is evidence that written AJTs can still be used for CS. Koronkiewicz and Ebert (2018) explicitly compared written and aural modalities for CS AJTs and found that only in the acceptable conditions was there a modality difference, with the aural modality improving the ratings.

The second factor is that CS is connoted very negatively in some communities and that this may influence ratings. Indeed, Badiola et al. (2018) found that attitudes *can and do* influence ratings, but only in one direction, similarly to what has been found for the influence of the written modality. The acceptable conditions were rated higher by participants with a more positive attitude towards CS, but for the unacceptable conditions, attitudes towards CS did not influence ratings.

AJTs have been extensively applied in code-switching research (e.g. Bandi-Rao and Dikken 2014; Cantone and MacSwan 2009; Koronkiewicz 2019a; Parafita Couto and Rodríguez-González 2019; Stadthagen-González et al. 2018; Van Dulm 2007; and many of the studies in López 2018), going back to the very beginning of research on the grammar of CS (Aguirre 1976; Timm 1975). Grabowski

(2011) showed that acceptability judgments of CS remain stable over time, which indicates reliability.

However, the use of AJTs in a multilingual setting *does* come with some additional complications. For example, Ebert and Koronkiewicz (2018) point out that it is important to include monolingual control items. These monolingual control items can be used to verify whether the participants in the AJT have the grammatical distinction under investigation as part of their monolingual grammar. If bilingual speakers do not discern a contrast in a monolingual setting, there is no reason they would detect that contrast in mixed-language sentences. For example, in chapter 8, the monolingual control items were used to exclude participants that did not negatively rate sentences such as (1), which monolinguals would rate as unacceptable.

(1) Danni cleans daily the kitchen.

Koronkiewicz (2019a) argues for the use of another type of control items in CS-mode as well. These items are CS control stimuli for which it has been well-established in the literature how their acceptability tends to be rated. Data from participants whose responses do not reflect the expected distinctions should be analysed separately. For more on the control items I used, I refer readers to appendix c.

The experiments described in chapter 8 include monolingual stimuli and CS control stimuli. However, as data collection for chapter 7 took place before the publication of Ebert and Koronkiewicz (2018) and Koronkiewicz (2019a), the survey that was used to collect the data for that chapter only includes target stimuli.

I will end this section with a note on the construction of the stimuli in this dissertation. Macaulay and Brice (1997) is a seminal study that investigated example sentences in syntax textbooks. It concludes that gender-bias and stereotyping is rampant in such example sentences and "that little has changed over the

past twenty-five years" (p 798). A more recent study investigated articles in three leading journals of the field and found that "this pattern has remained stable, with very little change, over the course of the past twenty years" (Kotek et al. 2020, p 514). Examples of the stereotyping and gender-bias found by these authors include male-gendered arguments being more frequent, more often standing for subjects/agents who are more often engaged in "intellectual activities", and who perpetrate violence more often, … than the referents of female-gendered arguments. I am convinced that these imbalances are the result of subconscious biases, so in the development of the stimuli, and the example sentences used throughout this dissertation, I made a conscious effort to avoid such gender-based stereotyping. Non-binary names used in examples in this dissertation (though none are included in the experimental stimuli to avoid introducing confounds) were sourced from Conrod (2020).

Call me an optimist if you will!

3.2 STATISTICS

The data collected for this dissertation and discussed in chapters, 5, 7 and 8 will be subjected to statistical analysis. Introductory textbooks to statistics will tell you that the goal of (inferential) statistics is using data from a (random) sample to draw conclusions about the population as a whole (Bolstad and Curran (2016, p 15); Desagulier (2017, p 151); Heumann et al. (2016, p 3); Mendenhall et al. 2012, p 3). In order to do so, statistical tests are required.

In this section, I outline what statistical methods I will use to analyse my data, and my motivations for doing so. The writing of this section has been greatly facilitated by the reading of Baayen (2008), Bross (2019a), Claes (2020), and Gries (2009).

3.2.1 Classic tests of significance

The fourth is the ratio scale.

There is some discussion in the literature on how exactly the results from AJTs should be analysed. To understand this debate, we need to talk about types of variable scales. Stevens (1946) distinguished four types of measurement scales, three of which are relevant for this dissertation: nominal, ordinal and interval scales. Nominal variables have values that cannot be ordered. Gender is a typical variable that produces nominal scale data. An ordinal variable has values that can be ordered, but the differences *between* the values cannot be compared in a numerical way. Ratings obtained through AJTs form a typical ordinal scale. The difference between "completely unacceptable" and "unacceptable" is not well-defined and may be (or probably is) different from the difference between "acceptable" and "completely acceptable". Interval variables, then, are variables whose values are not only ordered, but also numerically meaningful. Height, weight, age, ... are all examples of interval variables (Heumann et al. 2016, p 6–7).

While rating in an AJT results in ordinal data, in the literature, you will often see similar data analysed as interval scales. This is because, when it come to Likert-type data (Likert 1932), there is a common confusion between Likert scales (interval scale) and Likert items (ordinal scales) (Carifio and Rocco 2007, p 107). When you ask people to rate their agreement with statements on a scale of one to seven (or five, or eight, ...), this is a Likert response format. The statement plus the response forms a Likert item. A "Likert scale" only arises when there are multiple (6–8) Likert items that zoom in on the same underlying "attitude object", with both negative and positive evaluations of that object (Carifio and Rocco 2007, p 113). In table 3.1, I illustrate a Likert scale for "love of ice-cream".[2]

2 For the development of a proper Likert scale, more consideration should have been given to the statements than I have in this quick example.

To what extent do you (dis)agree with the following statements?	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree
1. I think about ice-cream regularly.	☐	☐	☐	☐	☐
2. I regularly eat ice-cream.	☐	☐	☐	☐	☐
3. I prefer ice-cream over other desserts.	☐	☐	☐	☐	☐
4. I prefer other desserts over ice-cream.	☐	☐	☐	☐	☐
5. I rarely eat ice-cream.	☐	☐	☐	☐	☐
6. I never think about ice-cream.	☐	☐	☐	☐	☐

Table 3.1: Likert scale consisting of 6 Likert items for the attitude object "love for ice cream"

This example should make clear that eliciting AJTs in a linguistic survey does not result in a Likert scale. Hence, data from AJTs are to be treated as Likert items. In principle, then, there are a lot of statistical tests that are ruled out when it comes to acceptability judgments. When it comes to measures of central tendency for example, the mean is mathematically not appropriate for ratings of Likert items, but mode or median are. However, Stevens himself noted (emphasis mine):

> As a matter of fact, most of the scales used widely and effectively by psychologists are ordinal scales. In the strictest propriety the ordinary statistics involving means and standard deviations ought not to be used with these scales [...] On the other hand, [...] *there can be invoked a kind of pragmatic sanction*: in numerous instances it leads to fruitful results.
>
> Stevens (1951, p 26)
> as cited in Bross (2019a, p 47)

So while some researchers insist on the use of non-parametric tests for the analysis of Likert items (such as χ^2-tests, see Boone and Boone 2012), others argue that parametric tests are so robust that they can be used for Likert items (Norman 2010). There are even a few studies which have compared the results of parametric and non-parametric tests applied to linguistic data from Likert items and found no difference in the results (De Winter and Dodou 2010; Endresen and Janda 2016). I will contribute to the conversation by analysing the results with both parametric

tests (t-test, ANOVA) and non-parametric tests (χ^2-test, Kruskal-Wallis test) and compare their results. Should they diverge, I will base my analysis on the non-parametric tests, as these are more conservative. If the results of the non-parametric and parametric tests converge, this could be further evidence for the "pragmatic sanction" invoked by Stevens.

3.2.2 Models

> A model is a simplification or approximation of reality and hence will not reflect all of reality [...] Box (1976) noted that "all models are wrong, but some are useful." While a model can never be "truth", a model might be ranked from very useful, to useful, to somewhat useful to, finally, essentially useless.
>
> — Burnham and Anderson (2002, p 20)

Besides the classic statistical tests of significance, statistical models are gaining popularity in the field. Gries (2009, p 253) defines a model as "the relationship between predictors – independent variables and their interactions – and one or more dependent variables". Models are used because they are more flexible, as they can account for a lot of variability in the data (Cunnings 2012, p 370). Models are well-suited for the analysis of linguistic experiments, as they allow for the inclusion of the effect of multiple independent variables on dependent variable(s). A *mixed* model is a model with fixed and random effects (Baayen et al. 2008, p 242). Mixed models "offer the advantages of being robust with respect to missing data, of allowing covariates to be taken into account, and of providing insight into the full structure of your data, including the random effects." (Baayen 2008, p 266)

More about fixed and random effects can be found below

Including random effects – often for item and participant – is crucial in the investigation of linguistic variables, as we are not interested in "experimental effects present only in the individuals who participated in the experiment, but rather in effects present in language users everywhere" (Baayen et al. 2008, p 390). Using

statistical models for inferences about the data involves two processes: model specification and model selection. Model specification involves deciding which candidate models to consider when selecting the best one (Burnham and Anderson 2002, p 18). Model selection consists in comparing the candidate models and selecting the one with the best fit. Let's first discuss some background notions, for readers who are unfamiliar with modelling. A linear model builds on linear regressions, which are represented by functions such as the one in 3.1.

$$f(x) = \alpha + \beta x \tag{3.1}$$

In this function, α expresses the intercept (point at which the line *intercepts* the y-axis) and β the slope of the line. The bigger the absolute value of β is, the steeper the line. A positive β indicates a line that goes up, while a negative β means a downward line (Bross 2019b, p 7).

Mixed models contain random effects, which can be modelled with either a random slope, random intercept, or both. For a model describing results from an AJT, a random intercept for participant would indicate that you allow for some participants being intrinsically higher raters than others, while including a random slope will allow for the possibility that participants may differ in their highest and lowest ratings. The same reasoning can be made for each individual item. This is why including random effects for item and participant is a good idea when modelling data of the type used in this dissertation (Baayen et al. 2008). While some authors recommend keeping the random effect structure as maximal as possible (Barr et al. 2013), Matuschek et al. (2017) argue that it is important not to include too many random effects as this will diminish the statistical power of the model. They argue that it is recommended that you only include the random effect structure which is supported by the data.

As for the fixed effects, these are the effects that are predicted to influence your dependent variable (i. e. they are the indepen-

dent variables). How do you decide which fixed effects to include in your model? There are two main strategies (though an intermediate strategy is possible): forward selection and backward selection. The former involves starting from the most simple model and adding effects only if their contribution significantly improves the model, i. e. if the model can account for more of the variance in your dataset. The latter consists in starting with the most complex model and systematically removing effects unless their removal significantly worsen the model. In this dissertation, I will be adopting the strategy of backward selection, as advocated for by Claes (2020).

A common metric for model selection is Akaike's Information Criterion (AIC). The better the model, the lower the AIC. It is important to note that the AIC is meaningless in isolation and should only be used to compare models. When the difference between the AICs of two models is between 0 and 2, there is hardly any difference between them in terms of how well they fit the data. If the difference is between 4-7, the model with the higher AIC has "considerably less evidence" to support it, while if the ΔAIC is larger than 10, there is "essentially no [evidence]" to support it (Burnham and Anderson 2002, p 70).[3] When sample sizes are (relatively) small, as in the data I'll be discussing, it is advisable to use the corrected AIC (or AIC_c) (Burnham and Anderson 2002, p 66).

Mixed models can be linear or ordinal. Again, there is debate about whether one *has* to use ordinal models when analysing Likert item data, or whether it is possible to analyse this type of data with linear models (which are parametric). Some authors just use linear mixed models without comment (Cunnings 2012). Others remark that linear models "work reasonably well on rating data, even though the rating data does not strictly meet all of the

3 An alternative measure to compare models in the Bayesian Information Criterion (BIC), which imposes a greater penalty for the number of parameters (Fabozzi et al. 2014, p 403), but Burnham and Anderson (2002) argue for the use of the AIC over the BIC (p 37).

assumptions of the regression." (Gibson et al. 2011, p 521). One simulation study found that using linear mixed models on Likert item data "does not increase the Type I error rate substantially" (Kizach 2014). Bross (2019b) says that using ordinal models is "more conservative" (p 25). As the literature arguing for the use of linear models is pretty scarce, I will choose the more conservative route and use ordinal mixed models to analyse my data.

Type I errors are false positives.

3.2.3 Effect size or p-values

Most people who have experience interpreting research based on statistical reports are familiar with the importance of p-values, but there has been a growing realisation in the last few decades that they are insufficient for the interpretation of (social science) data (Ferguson 2009, p 532 and references therein). Wilkinson and Task Force on Statistical Inference (1999) recommend the reporting of effect sizes with their confidence interval (CI) rather than p-values. Similarly, Sprouse and Almeida (2017) insist that classic statistical hypothesis testing obscures the real value of syntactic experiments: estimates of the effect size (p 27).

In the following chapters I will be reporting both p-values *and* effect sizes. In models, the regression coefficient (remember the β in function 3.1) can be used as "an easily interpretable effect size measure" (Nieminen et al. 2013, p 3). If the confidence interval doesn't include zero, the effect is considered to be significant.

3.3 SUMMARY

This dissertation reports on data collected through AJTs. While these are controversial in some corners of linguistics, they have been shown to be reliable and stable. I have also discussed the additional complications with using AJTs to investigate CS and

outlined some precautions that can be taken to counter-balance some of the complications.

The data will be analysed both with classic statistical tests and through the use of statistical models. Both p-values and effect sizes will be reported. While performing these classic tests of significance is redundant when modelling is applied, modelling is not yet a standard in the field. So I will include the more traditional analyses to facilitate comparison with other studies in the field.

These models are more informative and powerful!

For the classic statistical tests, both parametric and non-parametric tests will be performed to contribute to the debate as to which methods of analysis are appropriate for statistical analysis of AJT-data. As for the modelling, I will fit ordinal mixed models and will use the AIC_c to select between models. I will use R for all the analyses (R Core Team 2020). I have made all datasets and R scripts available on OSF. For a fun beginners' guide to R, I can recommend Phillips (2016). The R packages I used to analyse the data will be cited in the relevant sections.

Yes, fun ☺!

Part II

GENDER AGREEMENT

Gender is the most puzzling of the grammatical categories. This is the opening sentence of Corbett's seminal work on linguistic gender. It is certainly a vexing puzzle, with dozens upon dozens of articles devoted to solving it. The following chapters are an attempt at adding a few more pieces to the puzzle.

4 LEVEL: BEGINNER (MONOLINGUAL)

4.1 INTRODUCTION

Grammatical gender is a topic that has long intrigued linguists (see chapter 6 for some nicely phrased evidence). It is a property of nouns, which is characterised by agreement.[1] Bloomfield (1935, p 192) provides an early (modern) description of grammatical genders as "arbitrary classes, each of which demands different congruence-forms in certain kind of accompanying phrases". The observation that agreement is the hallmark of grammatical gender goes back to at least the fifth century BC (Ibrahim 1973, p 14). For an overview of the history of the study of grammatical gender, I refer the reader to Ibrahim (1973, chapter 2).

Corbett (1991) is arguably *the* seminal reference for grammatical gender. He divides the gender systems into roughly two categories: semantic and formal gender systems. While gender systems may be on either end of this spectrum, there are systems that rely on both semantic and formal criteria and any system may have any number of exceptions (p 8). In a semantic system,

1 While a formal distinction is often made between concord (DP internal agreement) and agreement (DP external agreement) (see for instance Klassen and Liceras 2017), here I use "agreement" in an informal manner. There is some discussion whether concord should/can be accounted for by Agree (see Carstens (2000) for arguments for and Norris (2014) for arguments against). I follow Carstens (2000) and others in assuming concord is obtained via Agree. Additionally, in the second language acquisition literature, there is also often a difference made between "gender assignment" (agreement between D and the noun) and "gender agreement" (agreement between other probes and the noun, see Pieters 2020 and references therein). I do not draw this distinction and reserve "gender assignment" for "how all nouns are given gender in a certain language" (Kramer 2015, p 249).

Semantic critereon	Gender	Examples	Gloss
male human or god	masculine	aaɳ	man
		civaɴ	Shiva
female human or god	feminine	peɳ	woman
		kaaḷi	Kali
other	neuter	maram	tree
		viiṭu	house

Table 4.1: Gender assignment in Tamil, adapted from Corbett (1991, table 2.1)

the meaning of a word determines its gender. While all gender systems are semantic to some extent, in languages with a strict semantic gender system, semantics is sufficient in determining the gender of nouns (p 8). One example of a strict semantic gender system is Tamil. The gender assignment rules of Tamil are given in table 4.1.

It is important here to pause for a small terminological discussion. We can see from table 4.1 that the masculine and feminine grammatical genders correspond to male and female referents. The reader is surely familiar with the distinction that has been made over the last few decades between sex – a biological property – on the one hand and gender – a social category – on the other hand. In the sociological/psychological literature "sex" has entirely been replaced by "gender" when describing differences in a population (Archer and Lloyd 1982, p 17). However, matters are slightly more complicated than just replacing "sex" by "gender", as is pointed out in Ackerman (2019). They distinguish three types of gender: grammatical, conceptual and biosocial gender. Biological sex overlaps with biosocial gender, which is "the multidimensional property of an individual as determined by their biology and cultural norms of identity expression" (p 9). Conceptual gender is the "gender that is expressed, inferred, and used

by a perceiver to classify a referent" (p 3). It is this conceptual gender then that is the basis of the semantic gender system in Tamil (and similar languages) and this is also the term I will be using throughout the rest of the dissertation to talk about the gender of human referents. "Male" and "female" will be used as shorthand for "referent of the masculine/feminine conceptual gender". Occasionally I will have examples with animal referents, in which case I will use "sex". When I use "gender" without being specific about whether I mean grammatical, biosocial or conceptual gender, I mean grammatical gender.

And then of course, there is the additional complication that, while in traditional Western society these concepts may have been approached that way, none of them are entirely binary. There has been a recent rise in linguists investigating the influence of trans and non-binary gender-identities on grammatical gender and non-binary language use. For examples I refer the reader to Ackerman (2019) and Conrod (2019) and references therein. For the purposes of this dissertation, I set aside this complication.

Conceptual gender also plays a role in formal gender systems. As mentioned earlier, all gender systems are semantic at their core and a common type of formal gender system, prevalent amongst the Indo-European languages, is based on the conceptual gender of the referent. Though this type of system is semantic in the sense that nouns denoting males go in the masculine category, while nouns denoting females go in the feminine category, for other nouns there is no semantic reason to put them in the class they belong to. Russian provides an example of such a system. In Russian, nouns that denote referents which lack a conceptual gender are seemingly randomly distributed over the three genders, represented in table 4.2 as "residue".

With the exception of human referents with conceptual gender, the grammatical gender of nouns in Russian is not semantically predictable. In contrast, it is easy to see what gender a word is

Grammatical gender	Criterion	Example	Gloss
masculine	male + residue	žurnal	magazine
feminine	female + residue	gazeta	newspaper
neuter	residue	pis'mo	letter

Table 4.2: Gender assignment in Russian, adapted from Corbett (1991, tables 3.1 and 3.2)

by looking at its declensional class, as is shown in (1). The gender assignment rules of Russian are morphological. For example, *pis'mo* ('letter') belongs to declensional class IV, as it ends in *-o* in the nominative singular, and hence, *pis'mo* is neuter (Corbett 1991, p 34–36).

(1) *Morphological assignment rules for grammatical gender in Russian*:
   1. Nouns of declensional class I are masculine.
   2. Nouns of declensional class II and III are feminine.
   3. Nouns of declensional class IV are neuter.

Other languages with a morphological gender system are the Bantu languages, where the nouns are divided into classes (i. e. genders) depending on their prefix.

For some languages, however, the gender assignment rules are completely opaque. This means that without looking at agreement with for example a determiner, predicting gender is not possible. This is the case for German and Dutch. However, some nominalisation suffixes can determine the gender of the noun they form, as demonstrated in the Dutch example in (2).

(2) a. vrij-heid
   free-dom[F]
   'freedom'

b. dans-er
dance-er[M]
'dancer'

This type of gender assignment system is also called derivational gender, since derivational morphemes can determine gender. Section 4.3 provides a more in-depth discussion of the gender systems of Dutch (and English), including how conceptual gender and grammatical gender interact.

In Generative syntax, gender is part of the set of ϕ-features, together with person and number. These ϕ-features are all involved in agreement (Luraghi and Parodi 2008, p 130). In this chapter, we will take a deep dive in the ways grammatical gender has been analysed within a Distributed Morphology (DM) framework. We will start in section 4.2 with a detailed discussion of Kramer (2015), as this is the first detailed, cross-linguistic perspective on gender in a DM framework. Note that she exclusively deals with grammatical gender systems based on conceptual gender with two or three genders. Since this dissertation also deals with such gender systems, Kramer (2015) provides an excellent starting point for the discussion. In section 4.3, I will discuss the gender systems of the languages relevant to chapter 5.

Before we move on, a brief note on the agreement process is required. In classic Minimalism, agreement is driven by uninterpretable features. Recall from section 1.2, that uninterpretable features need to be deleted by the end of the derivation. This can happen through agreement or movement, but here I will focus on agreement. In any agreement relationship, there is a probe and a goal. The probe is endowed with an uninterpretable feature and "probes down" into the structure it dominates to find a matching *interpretable* feature on the goal. Then the probe values its uninterpretable feature, after which it gets deleted (Chomsky 2001, p 5).

The probe and goal are also known as the agreement "target" and "controller" respectively. This is a bit unfortunate since "tar-

get" and "goal" may be synonyms colloquially, but opposites in this technical sense. I will stick to "probe" and "goal" in these chapters.

4.2 KRAMER (2015)

Kramer (2015) is an in-depth, cross-linguistic account of grammatical gender systems based on conceptual gender within a DM approach. I will illustrate it by means of Spanish, which she analyses in detail. To understand her account, I'll first present some background on Spanish grammatical gender and how it interacts with conceptual gender and the system of thematic vowels.

In Spanish, nouns are assigned one of two grammatical genders: masculine or feminine. For some nouns, grammatical gender matches conceptual gender (3)–(5). For some pairs of nouns, a change in thematic vowel (see below) changes both grammatical gender and conceptual gender of the referent (4). There are also nouns where this change in conceptual gender of the referent is accomplished only through agreement, but without change in form of the noun (5). Then, there are nouns that are invariant in grammatical gender, but can refer to referents of either conceptual gender (6) (Butt et al. 2019, pp 1-5).

(3) a. el yerno
'the son-in-law'
b. la nuera
'the daughter-in-law'

(4) a. el amigo
'the (male) friend'
b. la amiga
'the female friend'

(5) a. el soldado/colega/agente
'the (male) soldier/colleague/officer'
b. la soldado/colega/agente
'the female soldier/colleague/officer'

(6) a. la víctima
'the (male or female) victim'
b. el genio
'the (male or female) genius'

As indicated by the glosses in (4)–(5), the masculine variant often doubles as the unspecified variant, though there are also some cases of the feminine being the version which is the unmarked version (7) (Butt et al. 2019, p 2).

(7) a. la gallina
'the hen' or 'the chicken'
b. el gallo
'the rooster'

As we can see in the previous examples, grammatical gender in Spanish can sometimes be predicted from the final vowel of the stem, i.e. the thematic vowel. Thematic vowels are also known as inflectional classes or word markers and are sometimes treated as gender markers. However, many researchers treat them as being distinct from linguistic gender markers (Burkholder 2018; Harris 1996; Kramer 2015; Ritter 1993). The main reason for treating these as distinct phenomena is that gender is not entirely predictable from the thematic vowel. This is illustrated in table 4.3.

There are just a handful of feminine nouns that take theme vowel I, and a larger – but limited – list of masculine ones that take theme vowel II. So while theme vowels are not 100% predictive, they are often a pretty good indication of grammatical gender. For details on how these theme vowels are inserted in the syntax and how they interact with grammatical gender, I refer

Declension class	Theme vowel	Examples	Gloss	
I	-o	a. lí-o	muddle	M
		b. man-o	hand	F
II	-a	c. dí-a	day	M
		d. cán-a	grey hair	F
III	-e/∅	e. padr-e	father	M
		f. madr-e	mother	F
		g. lápiz-∅	pencil	M
		h. luz-∅	light	F

Table 4.3: Spanish declension classes, adapted from Bermúdez-Otero (2013, p 4)

to Kramer (2015, section 10.4.2). I will set aside this issue for the remainder of the chapter.

The interaction between grammatical gender and conceptual gender is important to Kramer. She makes a distinction between natural and arbitrary gender. The former is the convergence of the grammatical gender of a noun and the conceptual gender of its referent, while the latter is the grammatical gender of a noun with a referent that either has no conceptual gender (in the case of inanimate referents, for example) or has a conceptual gender that mismatches the grammatical gender of the noun.

At this point we need to take a little detour through two topics in the gender literature. The first one is the (un)interpretability of gender features. The second one is the locus of grammatical gender.

As regards the first one, there is disagreement in the literature whether or not grammatical gender on nouns is interpretable or not. Chomsky (2001, p 5) defines uninterpretable features as features that enter the derivation without values. For Chomsky, "uninterpretable features" are synonymous with "unvalued fea-

tures" (Pesetsky and Torrego 2007, p 266). Since nouns enter the derivation *with* a valued gender feature, gender on nouns must be an interpretable feature in the Chomskyan sense.

However, Pesetsky and Torrego (2007) propose that valuation and interpretability of features are independent concepts. For them, interpretable features are features which make a "semantic contribution to the interpretation of that item" (p 264) and it is *unvalued* features, rather than uninterpretable features, that cause a derivation crash if they are not deleted. If uninterpretable features are not deleted, they are simply ignored by the semantics, according to Pesetsky and Torrego (2007, p 290).

In this sense, it is less clear whether grammatical gender is interpretable or not. Carstens (2010) argues, on the basis of Romance and Bantu data, that there is no semantic contribution of gender features to the meaning of nouns. The fact that French *avion* (aeroplane) is masculine and *voiture* (car) is feminine makes no difference to the interpretation of these nouns. Hence, she concludes, grammatical gender is uninterpretable (p 35).

Others argue that grammatical gender *does* make a semantic contribution (Pesetsky and Torrego 2007, fn 31; Picallo 2005, 2008). They claim that grammatical gender "encode(s) mental processes of entity/object categorisation" (Picallo 2005, p 107).

Kramer (2015), who adopts Pesetsky and Torrego's distinction between (un)valued and (un)interpretable features, argues against an either fully interpretable or fully uninterpretable account of grammatical gender. The distinction she makes between arbitrary and natural gender maps onto a difference in interpretability: arbitrary gender is uninterpretable, while natural gender is interpretable. I will return to my own assumptions regarding this debate in section 4.4.

Let's move on to the second issue: the locus of the grammatical gender feature. This is a hotly debated topic in the literature. It seems that there are several possible options for the location of

Some papers that have in their title *a reference to the location of gender: Bobaljik (2006), Kramer (2016), Picallo (2017), and Ritter (1993)*

grammatical gender available in the literature. Gender is possibly a feature on the root, on little *n*, a dedicated functional projection of its own, or a post-syntactic operation. Kramer (2009) summarises the positions in the previous DM literature up until that point as follows:

> In the Distributed Morphology literature (see e.g., Marantz 1997, 2001, Arad 2003, 2005, Embick and Noyer 2007, Embick and Marantz 2008), the idea has been pursued that all lexical categories are made up of a category-neutral root and a category-determining head. [...] It is often assumed that roots have no syntactic or semantically active features, i.e., that they do not possess any features that drive syntactic operations or that are interpretable at LF. [...] However, they do have so-called diacritic features which encode root-particular quirks like inflection class [...] it could be that gender would also be a type of diacritic feature that can occur on a root. [...] To the best of my knowledge, it has not been specifically proposed in previous literature that gender is a diacritic feature on the root (at least insofar as gender has been treated as distinct from inflectional class). However, the idea that gender is a feature on *n*, the 'nominalizing' head, has received some support.
>
> Kramer (2009, pp 118ff)

For Amharic, Kramer (2009) concludes that a mixed approach is necessary, placing the gender not on either the categorising head (*n*), nor on the root itself, but on both these elements. She argues this is necessary to account for the differences between natural and arbitrary gender. Atkinson (2015) takes a similar approach for French. In Kramer (2015), the monograph partially based on the 2009 dissertation, there is a change in analysis.[2]

2 I would say that this change is a change toward the default position in the DM framework. One of the main tenets of DM is that roots are category-less, and as Acquaviva (2009) remarks "if a root has a feature that presupposes a category, then it is not really category-free" (p 2). This seems to be indeed a logical consequence and is in agreement with De Belder and Van Craenenbroek (2015, p 632) ("a root position is like a Bermuda Triangle for grammatical features") and seminal work by Harris (Harris 1996, fn 15: "Roots have no *mor-*

Both natural gender and arbitrary gender are on *n*, but the former is interpretable (*i*), while the latter is uninterpretable (*u*).

Kramer's analysis of Spanish relies on the fact that Spanish nouns can be derived by the Merger of a root with any of four different types of *n*s (cf example (11) in section 1.3). These *n*s are listed in (8). Merger with the plain *n* results in default (i. e. masculine) agreement.

(8) Types of *n* in Spanish (Kramer 2015, p 96)

a. *n i*[-F] → masculine natural gender
b. *n i*[+F] → feminine natural gender
c. *n* → no (or unknown/irrelevant) natural gender
d. *n u*[+F] → feminine arbitrary gender

Each root is subject to licensing conditions, which determine which of these roots can be selected by which of these *n*s. These conditions are listed in table 4.4. As can be seen, the roots in (3a), (3b), (6a) and (6b) can only be merged with one type of *n*, but the roots in (4), (5) and (7) can be merged with different *n*s. The difference between the roots in (4), (5) and (7) is that in (5), the words with opposing conceptual gender are root derived, while in (4) they are derived "via *-o/-a* alternation" and in cases like (7) via suppletion (Kramer 2015, p 95).

In this section I have laid out Kramer's (2015) analysis of grammatical gender. In the next section, I will discuss the relevant grammatical gender systems for chapter 5 and how these fit into Kramer's analysis. In section 4.4, I will explain which elements of her approach I adopt.

phosyntactic category, no gender, and no form of class affiliation"; emphasis mine)

		n *i*[+F]	*n* *i*[-F]	*n*	*n* *u*[+F]
$\sqrt{\text{YERN}}$	(3a)	×	✓	×	×
$\sqrt{\text{NUER}}$	(3b)	✓	×	×	×
$\sqrt{\text{AMIG}}$, $\sqrt{\text{AGENT}}$	(4), (5)	✓	✓	✓	×
$\sqrt{\text{VICTIM}}$	(6a)	×	×	×	✓
$\sqrt{\text{GENI}}$	(6b)	×	×	✓	×
$\sqrt{\text{GALL}}$	(7)	✓	✓	×	✓

Table 4.4: Licencing conditions in Spanish, adapted from Kramer (2015, table 6.2)

4.3 RELEVANT GENDER SYSTEMS

The gender systems of Dutch and French are historically related, as they are both Indo-European languages. Both of these are considered to be formal gender systems.

4.3.1 Romance: French

Most of the Latin neuter nouns were absorbed into the masculine.

While Latin had three grammatical genders (masculine, feminine and neuter), what remains in modern-day Romance languages such as French is a distinction between feminine and masculine (Maiden 2011, p 167f).[3] The grammatical gender of French nouns triggers agreement on determiners, adjectives, relative pronouns and participles. A summary of the relevant agreement patterns is shown in table 4.5.

Though the feminine reflex of adjectival agreement is orthographically represented with an *-e*, this difference is not always phonologically overt (9a). When it is overt, it may have a variety of effects. In Belgian French, where vowel length is phonemic,

3 There is a debate in the literature whether Romanian has two or three genders. For an overview, I refer to (Corbett 1991, p 150ff).

	-PL		+PL	
	-F	+F	-F	+F
indefinite article	un	une	des	
definite article	le	la	les	
adjectives	∅	-e	-s	-es

Table 4.5: Gender agreement in French

vowel lengthening may take place (9b). Another possible effect is making an underlying consonant overt (9c–d).

(9) a. M: acéré [asere] F: acérée [asere] (sharp)
b. M: cru [cʀʏ] F: crue [cʀʏː] (raw)
c. M: bon [bõ] F: bonne [bɔn] (good)
d. M: mauvais [mɔvɛ] F: mauvaise [mɔvɛz] (bad)

Similarly, the -s that orthographically represents the plural marker is rarely phonetically realised, though it does surface in cases of liaison.

While a combination of phonological and morphological rules allows for a certain predictability of gender assignment, there remains a substantial proportion of nouns – estimates range between 15.5% (Corbett 1991, p 60) and 20% (Lyster 2006, p 84) – for which the grammatical gender is unpredictable. Karmiloff-Smith (1981) showed that phonology plays a large role in early L_1 gender assignment to nonce words, but that this determining role of phonology decreases in older children. It seems that French does have phonological cues that are probabilistically linked to a specific gender, but (adult) speakers do not (and cannot) rely on these to assign gender to all nouns. Pérez-Pereira (1991) showed the same for Spanish.

As Spanish and French are closely related languages, it is unsurprising that the interaction between grammatical and con-

ceptual gender in French is quite similar to the one in Spanish. For a large number of nouns, there is a straightforward match between conceptual and grammatical gender (10) (compare with (3)).

(10) a. le garçon/mari/frère
the boy/husband/brother
b. la fille/femme/sœur
the girl/woman or wife/sister

There are also a large amount of nouns for which conceptual gender is morphologically reflected (11), with phonological effects similar to those explained for adjectival agreement (9). Note that the morphologically unmarked variant is not necessarily specified for conceptual gender of the referent (in the generic reading), while the marked variant always is. Sometimes the conceptual gender of the referent is *only* reflected in gender agreement and not morphologically marked on the noun itself, as is shown in (12) (Atkinson 2015, p 5–6) (compare with (5)).

(11) a. un étudiant
'the (male) student'
b. une étudiante
'the female student'

(12) a. le ministre
'the (male) minister'
b. la ministre
'the female minister'

Finally, the counterparts to (6) exist in French as well: the nouns that are traditionally known as epicene. For these epicenes, the conceptual gender of their referents does not interact with their grammatical gender, as shown in (13) (Atkinson 2015, p 7).

(13) a. le mannequin
'the (male or female) model'

b. la personne
the (male or female) person'

While Kramer does not treat French, it is easy to see how her analysis would account for the facts in French. Indeed, Atkinson (2015) provides an analysis of the French data within the approach proposed by Kramer (2009), and Fathi and Lowenstamm (2016) show briefly what Kramer's (2015) analysis would look like for the French data they discuss. In their paper, however, Fathi and Lowenstamm take an entirely different approach and they sum up the key differences between their approach and Kramer's as follows:

> Kramer views Spanish and (we must assume) French as systems implementing natural gender. For us, grammatical gender is the key to the system. Natural gender is but an epiphenomenon redundantly interpreting grammatical gender.
>
> Fathi and Lowenstamm (2016, p 506)

They start by stating that in French there are two types of nouns: one-gender and two-gender nouns. The former are nouns that exist in one version only: either feminine or masculine (14). The latter are pairs of nouns formed from the same root, but with differing grammatical genders (15). Note that these are not necessarily nouns referring to animates (15b) (Fathi and Lowenstamm 2016, p 478 and 484).

(14) a. le livre 'the book'
b. la maison 'the house'

(15) a. le chat – la chatte 'the (female) cat'
b. le grain – la graine 'the grain' – 'the seed'

In their approach, two-gender nouns may have a gender feature as part of their root: feminine two-gender nouns are formed with a complex root. For example, a root corresponding to the

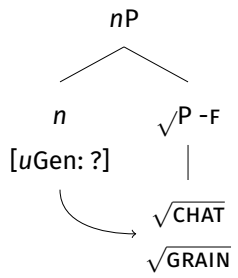
In Dutch, two-gender feminine nouns always have overt suffixes, see below.

lexical item $\sqrt{\text{CHA}}$ and a bound root[4] $\sqrt{+\text{F}}$ which houses the feminine feature. This bound root is spelt out as a CV-syllable. This accounts for the phonological alternation of two-gender nouns with a floating consonant (16). My summary glosses over some of the intricacies of the phonological alternation, but for details I refer to Fathi and Lowenstamm (2016).

(16) chat /ʃa/ 'cat'
chatte /ʃat/ 'female cat'

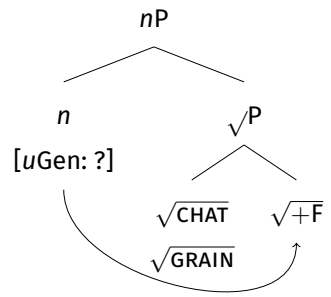
Nouns are derived by merging with *n*, which in their approach houses an unvalued gender feature. Valuation occurs when this feature on *n* probes down and finds either the lower $\sqrt{+\text{F}}$ (17b) (resulting in a feminine nominal) or doesn't (17a) (resulting in a masculine nominal). The complement of little *n* bears a value for gender in both (17a and b) "[b]ecause of contrastive marking" (p 494).

(17) a.



chat
grain

b.



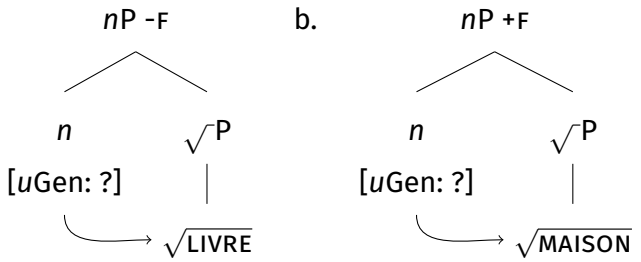
chatte
graine

For one-gender nouns, Fathi and Lowenstamm assume that neither $\sqrt{\text{LIVRE}}$ nor $\sqrt{\text{MAISON}}$ are merged with the $\sqrt{+\text{F}}$. The root phrase bears no value for gender, as $\sqrt{+\text{F}}$ is "absent from those representations". The probe on *n* will find no goal and this results

4 In traditional approaches, these would be affixes, but Fathi and Lowenstamm (2016) analyse affixes as roots too, albeit bound ones.

in “arbitrary gender”: gender is not predictable/compositional and needs to be learnt (p 497–499). They do not provide details on how exactly a specific root would get assigned a particular arbitrary gender, though presumably this could be worked out along the lines of Kramer’s licensing conditions (table 4.4).

(18) a. b.



While this approach seems like a promising avenue, I remain unconvinced by their argumentation for the difference between (17a) and (18), which is limited to “because of contrastive marking”. As this distinction is crucial for their approach, I am not inclined to adopt it. In addition, as Kramer (2016) points out, Fathi and Lowenstamm’s approach fails to provide an explanation for Kramer’s different-root nominals like *fille–garçon*. Since they are one-gender under the definition provided by Fathi and Lowenstamm (2016), their approach predicts that there should be words like *fille*, grammatically feminine, with exclusively male referents.

4.3.2 Germanic: (Belgian) Dutch and English

While there is some disagreement as to whether Proto-Indo-European had a two- or three-way gender system, it is generally agreed upon that (at least Late) Proto-Germanic had three grammatical genders: masculine, feminine and neuter (Schwink 2006, p 205). Modern Germanic languages have maintained this system to different extents; from losing grammatical gender altogether to maintaining the three-way gender distinction (Lopor-

caro and Paciaroni 2011, p 395). English belongs to the former end of the spectrum: it no longer has a system of grammatical gender, it is genderless. While it is true that pronouns do still present three gender distinctions (masculine: HE/HIM/HIS, feminine: SHE/HER/HERS and neuter: IT/ITS), it seems clear that pronominal choice is not a matter of grammatical agreement, but rather of semantic factors (Baron 1971; Vachek 1964).

Belgian Dutch (BD) represents the other end of the spectrum while Standard Dutch (SD) is an intermediate case; in SD the masculine and feminine distinction has been lost and it now only has a neuter/non-neuter (commonly known as "common") distinction.

Interestingly, while all varieties of Dutch maintain an active grammatical gender system, the pronominal system in both SD and BD is undergoing a shift similar to the one already completed in English: from one where pronominal choice is determined by grammatical agreement to one where semantic factors govern pronoun choice (Audring 2006; Audring and Booij 2009; De Vogelaer 2010).

De Vogelaer and De Sutter (2011) show that this semanticisation of pronominal choice is proceeding at a slower pace in BD compared to SD, and within Belgian varieties, the speed of change correlates with the degree to which gender agreement is overtly marked on adjectives and articles. In varieties where there is less overt gender marking (such as West Flemish), the change is quicker than in varieties where gender marking is morphologically more salient (such as East Flemish).

Let's return to grammatical gender. In the singular, SD has two genders: common and neuter (Broekhuis and Keizer 2012, p 6). In (19), the paradigm of definite and indefinite articles is shown. As we can see, in SD the gender difference is only marked in the paradigm of definite determiners.

(19)

a.	de	muur	c.	een	muur
	the[C]	wall		a	wall
b.	het	huis	d.	een	huis
	the[N]	house		a	house

As the focus of this paper is on BD, the gender system of SD will not be elaborated upon further, though of course it shows many similarities with that of BD. In contrast to SD, BD has three genders in the singular: masculine, feminine and neuter (Cornips and De Vogelaer 2009, p 1). This is demonstrated for agreement with the definite (20) and indefinite (21) determiners.

(20)

a.	den	bompa
	the[M]	grandpa
b.	de	bomma
	the[F]	grandma
c.	het	kind
	the[N]	child

(21)

a.	nen	bompa
	a[M]	grandpa
b.	een	bomma
	a[F]	grandma
c.	ee	kind
	a[N]	child

There are, however, some complications in the BD agreement paradigm. In certain phonological contexts, the difference between the masculine and feminine definite, and feminine and neuter indefinite determiners is covert. Both the masculine definite and neuter indefinite determiner end in an -**n** which is sensitive to phonological context and is deleted before all consonants except /h/, /b/, /d/ and /t/ (Taeldeman 1980, p 225). This means that the difference in the determiners in (21b) and (21c) is only overt because in the case of (21c) the -**n** has been deleted. For some nouns, agreement with the determiners does not show grammatical gender, as is illustrated in (22) and (23).

And yes, I also think it's weird that these consonants don't really form a natural class!

(22)

a.	de	vent
	the[M/F]	man
b.	de	vrouw
	the[M/F]	woman

(23)

a.	een	huisvrouw
	a[F/N]	housewife
b.	een	huis
	a[F/N]	house

This "phonological n" shows up all over the agreement paradigms in BD, as is shown in table 4.6. This table also shows that there is no gender agreement in the plural. The plural agreement forms are syncretic with the feminine singular ones.[5] Taking into account the phonological context in which this -**n** is deleted will be crucial in the development of the stimuli that are described in section 5.2.1.

	M	F	N	PL
definite determiner	de(n)	de	het	de
demonstrative	diene(n)	die	da	die
indefinite determiner	ne(n)	een	ee(n)	een
possessive pronoun 1SG	mijne(n)	mijn	mij(n)	mijn
adjectives	-e(n)	-e	∅	-e

Table 4.6: Overview of gender agreement in Belgian Dutch

Although BD may differ from French and Spanish in the number of grammatical genders it distinguishes, the interaction between conceptual and grammatical gender in BD follows a familiar pat-

5 Readers who are familiar with SD grammar may notice that the difference in definite/indefinite agreement of adjectives (shown in (i)) is not included in the overview.

(i)

adjectival agreement	C	N	PL
definite	-e	-e	-e
indefinite	-e	∅	-e

Taeldeman (1980) asserts that in Southern Dutch (i. e. Belgian Dutch), this "opposition has disappeared altogether" (p 225). The MAND-database (De Schutter 2005) however shows that there is variation between varieties spoken in Flemish Brabant as to whether or not this opposition is maintained (Vanden Wyngaerd 2012). Since this point is not crucial here, I will assume the sensitivity to definiteness is lost in BD and that the indefinite agreement forms are extended to the definite cases.

tern. For expository purposes in the following examples the determiners **ne(n)**, **een** and **het** indicate masculine, feminine and neuter respectively, though this is an oversimplification, as was discussed above. For animate nouns, there are a number of nouns where conceptual gender of the referent matches grammatical gender of the noun straightforwardly (24).

(24) a. ne(n) man/jongen/broer,…
a man/boy/brother…
b. een vrouw/meid/zus…
a woman/girl/sister/…

There are also nouns where there is a mismatch between conceptual gender of the referent and grammatical gender of the noun. This is most commonly due to the derivational diminutive suffix **-ke**, which always derives a neuter noun (25), though there are also some underived nouns with a mismatch between grammatical gender and conceptual gender/biological sex of the referent (26).

(25) het broer-ke/zus-ke, …
the brother-DIM/sister-DIM

(26) a. nen olifant
'a (male or female) elephant'
b. een muis
'a (male or female) mouse'
c. het paard
'the (male or female) horse'

In the following examples (27) the grammatical gender of the noun and conceptual gender of the referent also do not correspond ((27a) is uttered by a woman).

(27) a. Ik ben ook maar ne mens hè!
I am also but a[M] human INTERJ
'I'm only human!'

Van Dyck et al. (2009)

b. K=zit met ne vrouwelijke-n inspect-eur.
I=sit with a[M] female-M inspect-[AG.M]
'A female inspector is bothering me.'

De Schepper et al. (2019)

Kramer (2015) does not explicitly discuss (Belgian) Dutch, but her analysis (which she says is applicable to German) for the Papuan Lavukaleve is easily extended to account for the BD system. For three-gendered languages, she posits the following inventory of *n*s (28).

(28) Types of *n* in Lavukaleve (Kramer 2015, p 133)

a. *n i*[+F] → feminine natural gender
b. *n i*[-F] → masculine natural gender
c. *n* → neuter arbitrary gender
d. *n u*[+F] → feminine arbitrary gender
e. *n u*[-F] → masculine arbitrary gender

The important thing to note is that the plain *n* here is the neuter arbitrary gender, as this will come ~~back to haunt us~~ up in chapter 5. The licensing conditions for the roots that feature in the examples above can be found in table 4.7.

4.4 DISCUSSION AND CONCLUSION

While the account developed in Kramer (2015) does an excellent job of capturing the patterns she describes, there are some elements that leave me unsatisfied. For example, the analysis of different-root nominals (yerno–nuera, *garçon–fille*, **man–vrouw**) seems off. Under Kramer's analysis, these roots are identical but

		n *i*[+F]	*n* *i*[-F]	*n*	*n* *u*[+F]	*u*[-F]
$\sqrt{\text{MAN}}$	(24a)	×	✓	×	×	×
$\sqrt{\text{VROUW}}$	(24b)	✓	×	×	×	×
$\sqrt{\text{OLIFANT}}$	(26a)	×	✓	×	×	✓
$\sqrt{\text{MUIS}}$	(26b)	✓	×	×	✓	×
$\sqrt{\text{PAARD}}$	(26c)	×	×	✓	×	×
$\sqrt{\text{MENS}}$	(27a)	×	✓	×	×	✓

Table 4.7: Licencing conditions in BD

for their licensing conditions. Indeed, Kramer (2015, p 52) argues that "there is no inherent male-ness or female-ness to the roots themselves". She clearly feels that this is a feature, rather than a bug, continuing: "this approach has the added benefit of keeping the roots free from features that are associated with particular categories, like gender".

However, one does have to wonder about languages without grammatical gender that nevertheless do have pairs of roots as listed above. If it is desirable that roots contain no semantic information about (conceptual) gender, then where does this information go in, say, English? If the difference between *garçon* and *fille* is only captured by the type of *n* these roots merge with (*i*[-F] and *i*[+F] respectively), how does Kramer derive the difference in meaning between SISTER and BROTHER, if English has only one type of *n* available: the plain *n*? I think she would be forced to concede that this information is in the semantics of the roots $\sqrt{\text{BROTHER}}$ and $\sqrt{\text{SISTER}}$.

And what about languages such as Standard Dutch (and Swedish, and Danish) that are historically three-gender systems based on conceptual gender, but have lost the masculine/feminine distinction in the relatively recent past? In footnote 86, Kramer suggests that these languages are a potential candidate for a hypothetical language she describes where the agreement forms

of the masculine and feminine determiner are "accidentally homophonous" (2015, p 84). She seems to conclude this on the basis that in these languages the pronominal system still distinguishes male from female referents. However, as I've discussed in section 4.3.2 the pronominal system of SD (and to a lesser extent BD) is undergoing change and pronominal choice is mainly determined by semantic factors, rather than grammatical ones. It seems strange to argue then, as Kramer seems to do, that SD common nouns are vacuously marked for masculine or feminine gender.

Besides some animate nouns in languages without grammatical gender containing information about the conceptual gender of their referents, there are some other arguments to be offered for including information about conceptual gender in the meaning of the root. Consider the Dutch words in (29). If information about the conceptual gender of the referent is encoded on *n*, why does this information still seem to be present in the verb and adjective here? Even in items where the gendered component of the meaning has been lost (30), these still evoke their referent, gender and all.[6]

(29) a. ont-man-en
de-man-INF
'to emasculate/castrate'

b. ver-wijf-d
PREFIX-woman(pejorative)-ADJECTIVALISER
'feminine' (usually used for pejoratively for men)

6 The question of course is whether the words in (29) are root derived or word derived (i. e. derived from the noun, which contains an *n*). If they are word derived, Kramer could account for them the same way she does for German diminutives: the interpretable gender feature is present on a lower *n* head and is interpreted, since each merger with *n* triggers Spell-Out. Unfortunately, I have not found any good diagnostics to determine the (root vs word derived) status of the items in (29). Based on argumentation in Don (2005) I'd cautiously say they are root derived, though those in (30) clearly seem word derived.

(30) a. ouwe-hoer-en
old-whore-INF
'complain'

b. ver-broeder-en
PREFIX-brother-INF
'reconcile/fraternise'

A final argument here is an interesting minimal pair in BD (31). The word **mens** 'human being' has two possible grammatical genders: masculine or neuter. The masculine can refer to either humankind or an individual, in which case the individual must be male. The neuter on the other hand *must* refer to a female individual (pejoratively). Accounting for this minimal pair is only possible by assuming that the conceptual gender of the referent is part of the lexical meaning of the root, or in this case the meaning of the root in the context of certain *n*s. That is to say, the root $\sqrt{\text{MENS}}$ in the context of the plain *n* would be interpreted as referring to an annoying woman (31b), whereas this same root would be interpreted as either referring to a male individual (31a), or generically meaning 'human being' (recall (27a)).

(31) a. Die-ne mens heeft dat gedaan.
That-M person[M] has that done
'That guy/*woman/*person did that.'

b. Dat mens heeft dat gedaan.
That[N] person[N] has that done
'That awful woman/*man/*person did that.'

So while I adopt Kramer's approach to grammatical gender, I would like to propose that, for some roots (e.g. the roots in yerno–nuera, SISTER–BROTHER, *garçon–fille*, **man–vrouw**), conceptual gender of the referent, rather than being introduced by a type of *n*, is represented as semantic information in the lexical meaning of that root. Just like the root $\sqrt{\text{DOG}}$ is linked to an Encyclopedic entry in which information like 'pet', 'four-legged', and 'barks' is stored, a root like $\sqrt{\text{SISTER}}$ is is linked with an en-

try that includes information like 'female referent'. In contrast, the roots of words like **muis** or **olifant** (see (26a),(26b)) do not contain lexical information as to the gender of their referents.

One might argue that this leads to the loss of a powerful generalisation: when a part of the root's meaning is that it applies to a male referent, it gets masculine grammatical gender, and feminine gender when we have a female referent. However, as Kramer herself notes (scare quotes in original):[7]

> One possible solution would be to link up a "female" or "male" meaning component in the Encyclopedia entry for the root directly to the *n* choice, such that any "female root" is only interpretable under *n*[+FEM] and any "male root" under *n*[-FEM]
>
> Kramer (2015, p 52)

Though this was the approach taken in Kramer (2009)!

She believes "such an approach would be misguided" (p 52) because she does not want to have information pertaining to gender on the root. As I outlined above, because of the existence of non-gendered languages (and languages where the grammatical gender system no longer easily maps onto conceptual gender), it is clear that this information *can* be part of a root.

How about same-root nominals like $\sqrt{\text{AMIG}}$ then? I follow Fathi and Lowenstamm (2016) in assuming they are derived via affixation, but I'll leave open the question whether these affixes are also (bound) roots, as they argue.

Since I assume that information about the conceptual gender of the referent of words like *frère* and **vrouw** are part of the lexical meaning of their roots, I no longer need Kramer's distinction between interpretable and uninterpretable gender features (on *n)*. I will assume that all features on the noun are interpretable

7 The solution offered by Kramer seems to be a two-way street. We know from the linguistic relativity literature that the grammatical gender of an inanimate noun can influence its conceptual representation (see for example Phillips and Boroditsky 2003). Additionally, there is psycholinguistic evidence that L_2 learners transfer gender categories over from their L_1, even when the L_2 is genderless (Ganushchak et al. 2011). So because "male" and "female" are linked to masculine and feminine grammatical gender, inanimates with masculine/feminine grammatical gender are represented with male or female qualities.

and will treat “valued” and “interpretable” as equivalent, following Chomsky (2001).

In the next chapter, we will level up by taking a look at how the facts from the code-switching literature fit into the picture.

5 LEVEL: ADVANCED (MULTILINGUAL)

In this chapter, the contribution of code-switching (CS) data to the theory of grammatical gender will be discussed. Section 5.1 provides an overview of previous research on grammatical gender agreement in the CS literature. Section 5.1.1 discusses the CS literature approaching the topic of gender within a DM framework and section 5.1.2 discusses default gender in French and (Belgian) Dutch. Sections 5.2 and 5.3 describe the studies that I conducted on gender agreement in Dutch-English and English-French CS respectively. Section 5.4 summarises the results of those studies and outlines some suggestions for future research.

5.1 LITERATURE REVIEW

Linguists' fascination with grammatical gender does not stop at the monolingual literature and it is a well-researched topic in the CS literature as well. This is not surprising, considering that a lot of the CS literature focuses on English-Spanish CS and that Spanish has grammatical gender whereas English has not. So whenever an English noun is used in a Spanish sentence, the question of gender agreement arises.

Poplack et al. (1982) is an early study of grammatical gender in multilingual communities. The authors examined the genders of English loans in Puerto Rican Spanish (as spoken in the US) and Montreal French, both languages with two genders: masculine and feminine. They examined the following factors (Poplack et al. 1982, p 4–5):

- conceptual gender (in their terms: physiological sex) of the referent
- phonological shape of the loanwords
- gender of the translational equivalent (see below)
- association with the gender of a host language homophone
- association of a borrowed suffix with a host suffix requiring a certain gender
- default/unmarked gender

They found that the following factors are the most dominant: Conceptual gender, when present, overrides all other factors. Where phonologically based gender rules can be applied, they have a strong influence on the gender agreement. The gender of the translational equivalent of the noun was assigned to 60% (Montreal French) and to 85% (Puerto Rican Spanish) of the borrowed nouns.

While this was a study of loanwords, similar factors have been found in CS research on gender agreement. Most often, two strategies are observed in the literature on Spanish-English CS: analogical gender agreement and default gender agreement. In the former strategy the noun is assigned the gender of the translational equivalent equivalent of the noun (1a), while in the latter strategy, the default gender is chosen (1b). For Spanish, "[t]here is little doubt that masculine is the unmarked or default gender" (Harris 1991, p 43).

(1) ENGLISH-Spanish

a. la HOUSE
the[F] house
translational equivalent of HOUSE: casa[F]

b. el HOUSE
the[M] house
default/unmarked gender in Spanish: M

Parafita Couto (2019) presents an overview of the literature on code-switching across different language pairs and communities.

She remarks that the analogical gender agreement strategy is preferred by L_1 speakers of the gendered language (or speakers whose dominant language is the gendered one), while simultaneous bilinguals and L_2 speakers of the gendered language prefer the default agreement strategy (Balam et al. 2021; Jake et al. 2002; Klassen and Liceras 2017; Liceras et al. 2016, 2008; Valenzuela et al. 2012).

While most studies find a masculine-as-default strategy for CS with Spanish, Parafita Couto et al. (2015) found evidence for feminine as default in Spanish-Basque mixed determiner phrases (DPs). Basque nouns do not have gender, just like English nouns. Though this seems to provide evidence against masculine as the default gender in Spanish, it seems that phonological factors are at play.

The Basque post-nominal determiner *-a* is often borrowed together with the noun. This *-a* is then interpreted as a feminine word marker (see section 4.2), resulting in feminine gender agreement. This is illustrated in (2) (Parafita Couto et al. 2015, p 34).

(2) la *illar-a* *lodi-a* [Spanish-*Basque*]
the[F] pea-DET fat-DET
– Spanish equivalent of pea: guisante[M]
'the fat pea'

Even when the *-a* was not present the preference for feminine gender was observed, suggesting that the effect is extended beyond that phonological context. Badiola and Sande (2018) however found that the feminine default only arose in that particular phonological context and when the Basque determiner was not present, masculine gender was preferred for Basque nouns.

While the correlation of these different agreement strategies with different bilingual profiles is a robustly demonstrated phenomenon, community norms may play a role as well. For exam-

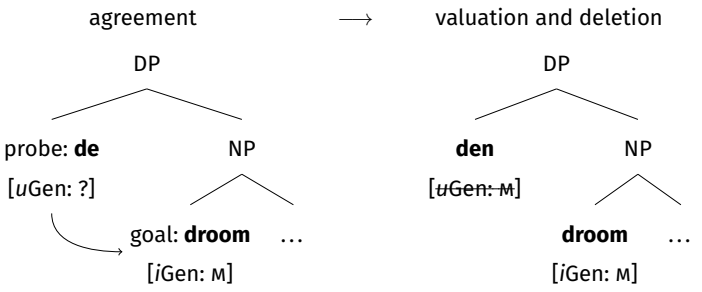
ple, although Poplack et al. (1982) report that the conceptual gender of the referent overrides all other considerations in gender agreement with loanwords, Balam (2016) reports that the tendency towards masculine default is so strong it even wins out over the conceptual gender of the referent, with striking examples of mixed DPs such as el VIRGIN MARY and un HOUSEWIFE (p 420) for Spanish-English CS in Northern Belize. Furthermore, Delgado (2018) reports a masculine default strategy overall, but with nouns of the familial sphere, gender of the translational equivalent (TE) is preferred by heritage speakers of (Mexican) Spanish in the "Chicagoland area". The importance of community norms was also found by Aaron (2015) in the gender agreement with loanwords.

Wikipedia tells me it is the nickname of the Chicago Metropolitan Area!

It's interesting to consider where these agreement strategies come from. Recall from section 4.1 that in any agreement relationship, there is a probe and a goal. The probe is endowed with an unvalued feature and probes down into the structure it dominates to find a matching *valued* feature on the goal. Then the probe values its unvalued feature, after which it gets deleted (Chomsky 2001, p 5).

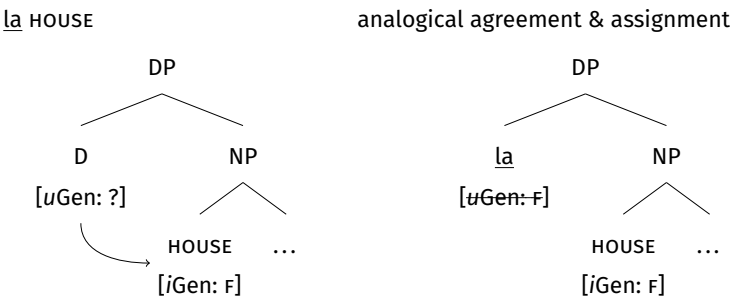
In (3), this process is illustrated (in a simplified manner) for the Dutch DP, **den droom**, 'the dream (masculine)'. The determiner **de** is endowed with an uninterpretable gender feature, which it needs to delete after valuation. To do so, it probes down into the structure it dominates and finds a goal with an interpretable gender feature, the determiner values its feature against the masculine gender feature of the noun and the uninterpretable feature is deleted.

(3)



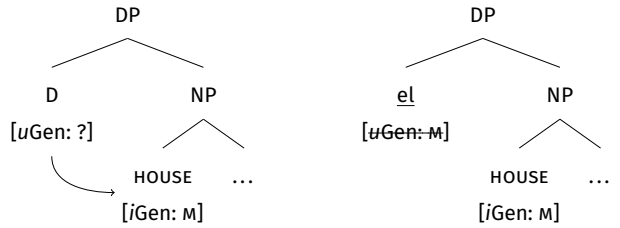
Let's return to the gender agreement strategies in CS. In (4), the analogical gender assignment strategy is shown. Here, we can see that feminine gender has been assigned to the English noun HOUSE.

(4)

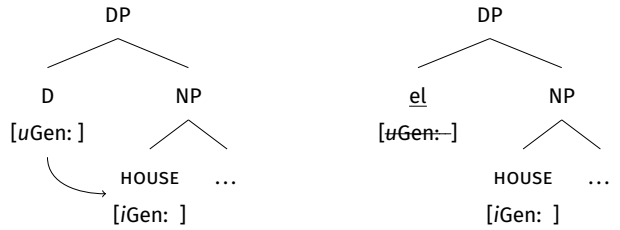


In (5), two options for the default strategy are shown. In (5a), there is also gender assignment. But instead of the analogical gender being assigned to the English noun HOUSE, the default gender is assigned. In (5b) on the other hand, no gender is assigned to the English noun and it remains genderless; it is not specified for gender. This latter analysis assumes that masculine gender agreement in Spanish it not just default, it is also unmarked.

(5) a. el HOUSE — default agreement & assignment



b. el HOUSE — default agreement, no assignment



The idea that default gender is actually unmarked gender is fairly common. Corbett and Fraser (1999, p 56) even remark that "*[n]aturally* the notion of default is connected to markedness [emphasis mine]". For Spanish, there is a fair amount of literature arguing for the unmarked status of the masculine (Harris 1991; Kramer 2015; Picallo 2005, 2008).[1]

As a CS-researcher, it is this latter option (the one represented in (5b)) that is more attractive to me, as it would mean that early (for example) Spanish-English bilinguals don't assign a gender to each English noun that occurs in a construction where agreement with Spanish probes is necessary, but rather that they keep each English noun "as is".

In the option represented by (5a), there is no real difference between the analogical and default strategy. In both cases, nouns get assigned a gender. In concrete terms, both early (or English-dominant) and late (or Spanish-dominant) bilinguals treat English nouns inserted into Spanish the same, only they differ in

1 There is also neurocognitive evidence that the masculine and feminine gender (in Spanish) are qualitatively different, as they are processed differently (Beatty-Martínez et al. 2020).

the gender they assign to English nouns; these get assigned masculine (extremely consistently sometimes as reported in Balam 2016) for some speakers and masculine or feminine (depending on the analogical gender) for others. If you assume that the default gender is the unmarked one however (cf Kramer's plain *n*), the default agreement strategy of early bilinguals follows automatically and gender assignment only needs to be assumed for late (or Spanish-dominant) bilinguals.[2]

I have the impression that most of the CS literature agrees, as they account for default gender agreement as in (5b) (Delgado 2018; Flores 2018; Liceras et al. 2008), though not all authors are explicit about it. Burkholder (2018) and López (2020) are exceptions, as they assume that the masculine determiner in Spanish is marked for gender.

Another phenomenon consistently reported in the literature (Parafita Couto 2019) and which will prove relevant in the next section is an asymmetry in mixed DPs. In mixed-language corpora, it is often just one of the two languages that provides the determiner in mixed DPs. This asymmetry has been observed in corpora of Spanish-English CS (Blokzijl et al. 2017; Jake et al. 2005), Welsh-English CS (Deuchar 2006) and German-English CS (Jorschick et al. 2011). All these bilingual corpora contain CS between English and a language with grammatical gender. It is the case that the language *with* grammatical gender overwhelmingly provides the determiner in mixed DPs in all of these corpora, *except* the one in Blokzijl et al. (2017), where it is English that provides the determiner.

2 One may raise the objection to (5b) that this would mean that the uninterpretable feature on the goal (the determiner) would fail to be deleted, as it cannot find a goal with a gender feature to be valued. However, I follow Chomsky (2000a, p 124) and others in assuming that a probe values and deletes the full set of ϕ-features in "one-fell swoop".

5.1.1 Code-switching and DM

There are (to my knowledge) less than a dozen papers that include a discussion of grammatical gender agreement in CS in a DM (or similar) framework, the majority of which deal with English-American Norwegian CS (Alexiadou et al. 2015; Grimstad et al. 2014, 2018; Riksem 2017, 2018; Riksem et al. 2019).[3] Most of these papers however have as their primary goal arguing that exoskelatal models of grammar (such as DM) are more suited to analyse CS data than lexicalist frameworks. They use gender agreement as one of their arguments in favour of exoskeletal approaches, so they only sketch their approach in broad strokes, and they don't provide an in-depth investigation of grammatical gender.

As these papers have several authors in common, it is not surprising that the analyses presented in them are very similar. They put grammatical gender in a functional projection in the nominal domain usually designated by either "gender phrase" (see for example Grimstad et al. 2014) or generically as "functional projection"/"F", in which case it also houses other ϕ-features (see for example Grimstad et al. 2018; Riksem 2018). While these papers provide valuable arguments for the use of an exoskeletal approach to CS, the analyses for gender are not detailed enough to be of further use in this chapter. Note that Kramer (2015, section 2.3.1) argues against locating grammatical gender in a gender phrase or other functional projection on the basis of her Amharic data (p 25).

Burkholder (2018) provides a more detailed analysis and an experimental investigation of grammatical gender in English-French CS within DM. While the new data that are discussed are interesting and a great addition to the field, I think the analysis presented is flawed. For a complete overview and thorough

3 Liceras et al. (2008) also explicitly adopt DM, but their analysis does not discuss what happens below the level of N.

	Kramer (2015, p 96)	Burkholder (2018, p 8)
the	[DEF][5]	[DEF]
el	[DEF], [-PL]	[DEF], [-F]
la	[DEF], [+F], [-PL]	[DEF], [+F]
los	[DEF], [+PL]	[DEF], [+PL]
las	[DEF], [+F], [+PL]	[DEF], [+F], [+PL]

Table 5.1: Feature specification of definite determiners in Spanish & English

understanding of Burkholder (2018), I recommend reading the original paper. I will provide a brief sketch of the crucial elements needed to understand the substance of my criticism.

Burkholder's main goal is to provide an analysis of the mixed-DP asymmetry. While previous analyses (such as Liceras et al. 2008; MacSwan 2005) account for this in terms of the difference in feature specification of the determiners in Spanish and English, Burkholder argues that this asymmetry can be attributed to the relationship between Spanish roots and declension class.

Her main assumptions about grammatical gender mainly follow Kramer (2015), which was explained in detail in section 4.2, but she makes one change to Kramer's analysis that I would like to point out before continuing the summary of her approach. Let's take a look at the feature specification for the Spanish and English determiners according to Kramer and Burkholder (table 5.1).[4]

Whereas the Spanish determiner el is not marked for gender in Kramer's approach in Burkholder's it *is* specified for [-F]. Remember that nouns are formed by the Merger of a root with a

4 I've used binary features as in Kramer (2015) and converted Burkholder's to match those.

5 There is no explicit mention of the feature specification of THE in Kramer, but I assume this is what she would say.

categorising *n*. In (6), I have repeated the inventory of *n*s in Spanish (first mentioned in (8) in section 4.2). Nouns with masculine gender can be derived by Merger of a root with either the plain *n* or the *n* which is marked for interpretable masculine gender (*i*[-F]).

(6) Types of *n* in Spanish (Kramer 2015, p 96)

a. *n i*[-F] → masculine natural gender
b. *n i*[+F] → feminine natural gender
c. *n* → no (or unknown/irrelevant) natural gender
d. *n u*[+F] → feminine arbitrary gender

The reason why both nouns with natural masculine gender (derived with an *n* marked with *i*[-F]) and arbitrary masculine gender (derived with the plain *n* unmarked for gender) end up with el is due to the subset principle. The determiner el wins competition for insertion despite being un(der)specified for gender. All other (Spanish) candidates for insertion are either overspecified or have an incorrect specification.

Burkholder abstracts away from natural gender and postulates just two possible *n*s (7) that can derive Spanish nouns. Note that there is no plain *n* for Burkholder.

(7) Types of *n* in Spanish (Burkholder 2018, p 8)

a. *n u*[-F] → masculine gender
b. *n u*[+F] → feminine gender

In a nutshell, Burkholder argues that the source of the mixed-DP asymmetry is that an English root can be selected by either the Spanish (gendered) *n* or the English (plain) *n*, while due to phonological constraints, Spanish roots can *only* be inserted in the context of declension class markers, which are tied to the Spanish *n*.[6] When an English root merges with an English *n*,

6 This is a simplification; for a full explanation I refer to pages 5–8 in the original paper.

and this noun merges with D, a regular unmixed DP will arise, as shown in (8).

(8) the house

In DPs such as those in (1) however, the English root merges with a Spanish *n*. This Spanish *n* is endowed with grammatical gender. Therefore the determiner will also have a gender feature, as it agrees with the noun. When Vocabulary Insertion (VI) occurs, a Spanish determiner (which is specified for gender) will win the competition for insertion over the English one (unspecified for gender), due to the subset principle.

Because Spanish roots, due to phonological constraints, cannot be merged with English *n*s, DPs such as THE casa are ungrammatical. Crucially, – and contrary to the previous analyses – Burkholder's analysis predicts that for a language with no inflectional classes/thematic vowels, no such asymmetry between English and that language will be found. She goes on to test that prediction for English-French CS, using both grammaticality judgments and a self-paced reading task. Her findings corroborate the predictions of her analysis.

There is, however, a critical issue with the analysis. Consider a mixed Spanish-ENGLISH DP such as the one in (9), which is attested in the Bangor Miami corpus (Blokzijl et al. 2017, p 7).

(9) a. los DORMS

b. [$_{\mathrm{DP}}$ los [$_{\mathrm{NumP}}$ [$_{n\mathrm{P}}$ *n* [$\sqrt{\text{DORM}}$]] -s]]

While Burkholder doesn't treat plural DPs, we could extend her analysis and conclude that in (9) we have an English root, merged with a Spanish *n*, merged with an English plural morpheme, merged with a Spanish determiner. The simple DP above would contain four switch sites, three of which are covert. While it is difficult to show that this is *not* the case, these different covert switch sites would be expected to correspond to patterns or similar asymmetries in the attested data. However, I am not aware

of any such patterns, nor does the original paper discuss any reasons whether these covert switch sites are independently motivated.

The lack of a treatment of plural DPs results in yet another issue. As I mentioned earlier, Burkholder's account for the DP asymmetry hinges on the fact that el is a better candidate for insertion than THE in mixed DPs such as (10).

(10) el BOY

The Spanish *n* that merges with the English root is equipped with a gender feature. After agreement, D will also be equipped with a gender feature, making el a better candidate for insertion. In Kramer's account, however, el is unspecified for gender, just like the English determiner THE. The problem for Burkholder's account is that a structure such as (11) is predicted not to arise due to the same principles of insertion. In the DP in this sentence, the determiner is not equipped with a gender feature if the root merges with an English plain *n* (similar to (8)), but it will be equipped with a plural feature. However, as is shown in table 5.1, Burkholder assumes English THE to be unspecified for number, making los a better candidate for insertion. I know of no literature asserting that structures paralleling (11) are not attested.

(11) Me gustan [$_{\text{DP}}$ THE HOUSES].
I like …

Even if we suppose that the English definite determiner *is* in fact specified for plural, and there is syncretism between the singular and plural form, we run into trouble. Note that in table 5.1, the plural masculine determiner los is unspecified for gender, just like in Kramer's account. This means that for Burkholder both *los* and plural *the* would have the same features and the DP asymmetry should then only be attested in the singular. I am

unaware of any studies showing this.[7] All in all, Burkholder's approach is interesting, but ultimately it fails to account for the data.

Another relatively in-depth analysis of grammatical gender within a DM framework is López (2020). In his approach, Spanish has two types of *n*: *n*[+F] and *n*[-F], while English just has the plain *n*.[8] In a DP like (12a), the English root $\sqrt{\text{TABLE}}$ has been selected by *n*[+F] by analogy with the selection that occurs for the Spanish root $\sqrt{\text{MESA}}$ (p 96). For (12b), the $\sqrt{\text{TABLE}}$ has been selected by the plain *n* as would happen in monolingual English. The probe (el) remains unvalued, since it does not find any gender features on the goal and "adopts default values" (López 2007, p 42), i. e. masculine.

(12) a. la TABLE
b. el TABLE

López's account offers some interesting suggestions. He argues that nouns in different languages have different "morphosyntactic frames" in which any type of root can be merged. The reason for the analogical gender agreement strategy is because there is a strong link between the roots of, for example, CHEESE and *fromage*. He goes so far as to argue that CHEESE and *fromage* are different exponents of the same root and that any difference in interpretation is taken care of by the interaction of pragmatics and the Encyclopedia, though he concedes that it is possible that the translational equivalents can also be different roots that just have significant overlap in the Encyclopedia (p 51–53).

This notion is interesting, especially if one considers code-switching between two related varieties such as BD and SD. Consider the pairs of words in (13). There is, as far as I can tell, no

7 Admittedly, I am also unaware of any (corpus) studies explicitly investigating plural DPs, but it seems like something people would remark on.

8 The reason why he assumes the masculine is marked [-F], rather than unmarked has to do with data from German-English code-switching.

difference at all in meaning between the SD and BD variants, despite the difference in pronunciation for the pair in (13b).

(13) a. het boek SD
den boek BD
b. de machine SD
het machien BD

The difference in gender agreement preferences for different bilingual types can be explained easily. L_1 speakers of Spanish or Spanish-dominant bilinguals would have a preference for inserting all roots (including the ones that are exponed by an "English" vocabulary item) in a "Spanish" morphosyntactic frame. This would result in the analogical gender agreement strategy. Early bilinguals (or English-dominant bilinguals) would prefer to use an "English" morphosyntactic frame for "English" roots, ending up with an *n*P without grammatical gender, accounting for the (overwhelming) default agreement strategy. This is a flexible system and can account for the variation that is found within communities as well. If the norm in a certain community is that certain "English" words have a grammatical gender (like the words in the familial sphere as described by Delgado 2018) even early bilinguals will merge such roots in a "Spanish" syntactic frame.[9]

It is important to note that this alternative approach fails to account for the mixed-DP asymmetry, which is consistently found in CS corpora (see section 5.1). However, as Burkholder herself notes, this asymmetry has not (yet?) been found in grammaticality judgment tasks (Liceras et al. 2008).[10] It also does not

9 I've put the names of the languages between scare quotes here, because I don't want to suggest that these roots are somehow marked for language membership, but rather that these are traditionally considered as belonging to these languages.

10 Burkholder argues that this is not problematic for her data as the "experimental task used [...] differs from those of previous studies in several ways. [...] As such, they [the participants] might not have the same opportunity to decon-

follow the same pattern for *all* CS communities. Blokzijl et al. (2017) found a mixed-DP asymmetry for their Spanish-English CS corpus, but instead of the language with grammatical gender providing the determiner, it is the genderless language that does. Whether or not the mixed DP asymmetry arises from the architecture of grammar (MacSwan 2005), processing factors (Liceras et al. 2008) or extralinguistic factors (Blokzijl et al. 2017) remains an open question. Consequently, it also remains an open question whether or not a purely grammatical account of gender and code-switching should be able to explain the mixed-determiner asymmetry.

Before we can go on to the newly collected data, we need to talk about which gender is the default gender in French and Dutch. We turn to this question in the next section.

5.1.2 Defaults in French and Belgian Dutch

For French, it is pretty uncontroversial that masculine is the default gender. It is the one which is used when the conceptual gender of the referent is unknown (14a), when the goal is a coordination of nouns with different genders (14b), and it attracts over 80% borrowings (Roché 1992, p 118). French also has slightly more masculine (56-57%) than feminine nouns (Becker and Dow under review, p 4; Roché 1992, p 113)

(14) a. Qui est heureux ?
who is happy[M]
'Who is happy?'

Schane (1970, p 293)

struct and consciously analyze their grammatical structure, thus relying more on their intuition as opposed to their metalinguistic knowledge" (Burkholder 2018, p 23).

b. Le garçon et la fille sont petit-s.
the[M] boy and the[F] girl are small[M]-PL
'The boy and the girl are small.'

Schane (1970, p 291)

In contrast to what she assumes for Spanish!

The question is, is it also the *unmarked* gender? Schane (1970, p 289) argues that this is the case on a morphological basis. Burkholder (2018) also assumes so, and while she doesn't provide evidence, it is easy to see that masculine agreement is chosen when the goal is not marked for gender. In (15), the probe *honteux* 'shameful' displays masculine agreement, while the goal *mentir* 'to lie' is an infinitive, and is presumably not marked for gender.

(15) Ment-ir est honteux
lie-INF is shameful[M]
'Lying is shameful'

Roché (1992, p 114), translation mine

The case for default in Belgian Dutch (BD) is somewhat more complicated.[11] Because there is no gender agreement in the plural, and there is no agreement on predicative adjectives, tests such as (14) cannot be applied. Most evidence however, seems to point toward the *neuter* as being the unmarked gender.

Rooryck (2003) develops an analysis on the basis of morphology. His account involves two levels of specification, the gender and subgender level. He argues that neuter is completely unmarked for gender, while masculine is marked for gender, but unmarked for subgender (16).

11 I am leaving Standard Dutch (SD) out of consideration here, though some of the facts I discuss apply to SD as well.

(16) Rooryck (2003, p 5)

a. feminine: [Gen: [Subgen: fem]
b. masculine: [Gen: [Subgen:]
c. neuter: [Gen:]

This absence of gender marking for neuter nouns would track well with the tendency for neuter agreement for non-nominal elements, which are unspecified for gender. Look at what happens when a determiner agrees with a verb (17). The determiner takes neuter gender, even though there is presumably no gender feature on the verb.

(17) Ik word hees van dat roepen.
I become hoarse of that[N] shout-INF
'That shouting is making me hoarse.'

One might ask if it is true that we are dealing with a verb here, and not with a noun that has undergone conversion. Compare (17), with the converted noun in (18). Here, the noun **geroep** (derived from the past participle) takes neuter gender as well.

(18) Ik word hees van dat ge-roep.
I become hoarse of that[N] PTCP-shout
'That shouting is making me hoarse.'

The difference between (17) and (18) becomes apparent when we try modification with one of the few adjectives[12] in Dutch that can only modify verbs and clauses and cannot modify nouns, such as **altijd** 'always' (Broekhuis 2013, p 465f). In (19a), we can see that **altijd** cannot modify the converted **geroep** while it can modify the unambiguous verb in (19b) and the verb preceded by a determiner in (19c).

12 Dutch does not have a grammatical class of adverbs, see section 8.2 for argumentation.

(19) a. *Ik word hees van dat altijd ge-roep.
I become hoarse of that[N] always PTCP-shout
intended: 'Always shouting is making me hoarse.'

b. Ik word hees van altijd roep-en.
I become hoarse of always shout-INF
'Always shouting makes me hoarse.'

c. Ik word hees van dat altijd roep-en.
I become hoarse of that[N] always shout-INF
'The constant shouting is making me hoarse.'

More generally, it seems that clauses, another type of non-nominal element, also take neuter agreement in determiners (20).

(20) Het bedenk-en van goede voorbeelden is moeilijk.
the[N] come.up-INF of good examples is hard
'Coming up with good examples is hard.'

On the basis of first language acquisition data, Roodenburg and Hulk (2009) propose that the neuter pronoun in Dutch is unmarked for not only gender but also number. This proposal is along the lines of Picallo's analysis of the Spanish pronoun lo (2005; 2008). However, despite the fact that examples (17)–(20) would seem to indicate that the neuter determiner is unmarked for gender too, Roodenburg and Hulk are less convinced about the unmarked-ness of the neuter determiner. Yet, Kramer (2015) proposes that neuter nouns are derived by Merger of a root with a plain *n* for languages such as Belgian Dutch, as was described in section 4.3.2.

So what can we predict about CS? And how can data from CS help us to better understand the gender systems of French and (Belgian) Dutch? For English-French CS the prediction is pretty straightforward, as the gender system of French seems to be understood pretty well. Early bilinguals are predicted to use masculine gender agreement for English nouns, while late bilinguals may prefer an analogical gender agreement strategy. This is indeed what has been found by Burkholder (2018).

For BD it seems that neuter is the unmarked gender. However, take a look at (21), where there are some utterances I informally collected from Facebook. Here the English loanwords (I would not classify these as code-switches, though it can't be excluded) are assigned masculine gender. While the posters here are L_2 speakers of English, these words don't really have a translational equivalent, so analogical gender agreement seems excluded here. Instead of taking neuter, they take masculine.

*Maybe **afspraakje**[N] for DATE, but these are not quite equivalent.*

(21) a. **precies nen** DATE
looks.like a[M] date
'looks like a date'

b. **wie neemt er ne** ONESIE **mee**?
who takes EXPL a[M] onesie with
'who is bringing a onesie?'

Similarly, the picture in figure 5.1 shows an ad for a local internet company to indicate that a hotspot is available. For all three nouns in that figure and in (21), neuter gender seems completely ruled out to me as a sequential Dutch-English bilingual.

Figure 5.1: Advertisement in Flanders

At this point it is useful to bring up the distinction that has long been made in the literature between normal case default and exceptional case default, going back to Pāṇini (Fraser and Corbett 1997, p 43). In the lexicalist frameworks where this distinction is used, the normal case default is the default that is assigned when the lexical entry is un(der)specified for grammatical gender, whereas the exceptional case default is assigned

"when the normal system breaks down", for example when the goal is non-nominal or when "some idiosyncratic feature of a lexical entry gets in the way of normal class assignment" (p 44). So the data we have discussed above for Dutch seem to identify the exceptional case default, rather than the normal case default. Unfortunately, tests for the normal case default are hard to apply to Dutch due to the limited gender marking and lack of gender distinctions in the plural in (Belgian) Dutch.

Capturing the distinction between the normal and exceptional case default in a DM framework is tricky as they "do not map onto useful distinctions" (Kramer 2015, p 141). However, based on a sample of languages, Kramer makes the following generalisations for three-gender languages:

(22) Default gender generalisations for three-gender languages (Kramer 2015, p 145)

a. Default gender agreement is neuter for all non-sex-differentiable controllers [i. e. goals]
b. Default gender agreement is not neuter for all sex-differentiable controllers [i. e. goals]

While she admits that (22b) is not without exceptions, she claims (22a) "holds absolutely" (p 146). This means that if we follow Kramer, against my intuitions and informally collected data, we expect neuter gender agreement to occur with inanimate code-switched nouns. So the examples in (21) constitute the first possible problem with her account. The second problem is that it is unclear how (22) is practically implemented in her analysis. Recall that in three-gender languages the plain *n* is characterised by neuter agreement. But how does she explain the use of another type of default? The other *n*s in her inventory are all equally specified. I will return to this issue in the sections that discusses the results of the survey that I now turn to (section 5.2.4).

5.2 TESTING DUTCH-ENGLISH CS

Considering sections 5.1 and 5.1.2, there are several questions we can ask when it comes to gender agreement to English nouns in Dutch-English CS.

1. Is there evidence for the mixed-DP asymmetry in Dutch-English CS?
2. What gender agreement strategies exist?
3. Is there a difference in gender agreement strategies across groups of bilinguals?
4. If a default gender strategy is used, what is the default gender in Belgian Dutch?

To attempt to try and answer these questions, I developed an acceptability judgment task (AJT).

5.2.1 Materials

For this survey, 36 preterite sentences were constructed. The preterite was chosen to avoid the simple present, as this tense has very different uses in Dutch and English. All the sentences had a ditransitive verb with a single noun insertion. All inserted nouns were preceded by a definite determiner and the DP containing them occurred in sentence-final position. Close cognates, both phonetically and orthographically (such as **vinger** – FINGER; **baard** – BEARD) and idiomatic expressions were avoided. An effort was made to make sentences sound as natural as possible given the constraints.

When the Dutch singular determiner was masculine or feminine, the noun began with a /h/, /n/, /d/, /t/ or a vowel, as these are the phonological contexts in which the /n/ of the masculine is *not* deleted and there is no syncretism between the masculine and feminine singular determiner. An effort was made to make sentences sound as natural as possible given the constraints.

The factors that were included in the stimulus-design were:

- language of the embedded noun (which was of course the opposite of the rest of the clause)
- number of the embedded noun
- when the determiner was Dutch singular:
  - gender of the Dutch determiner: M, F or N
  - gender of the translational equivalent of the noun: M, F or N

This resulted in a total of 12 conditions, which are represented in table 5.2. As we can see in the column headed by "n^{o}", the first 9 were presented in two lexicalisations (2*9=18), while conditions 10–12 were presented in six lexicalisations (6*3=18) for a total of 36 (18+18) sentences. The gender of the determiner and the TE of the noun were not a factor in conditions 10–12, as the English determiners and Dutch plural determiner and nouns do not have gender. In (23), some sample stimuli are presented. The full set of stimuli can be found in appendix b.

(23) a. **Mijn collega kocht den** CHAIR.
my colleague bought the[M]
condition 1

b. **Haar ouders betaalden het** RENT.
her parents paid the[N]
condition 8

5.2.2 Procedure

The survey started with a short background questionnaire. Some participants were at this stage screened out automatically for a variety of reasons (self-reported low proficiency/low use of either language, late acquisition of Dutch, ...). A short proficiency test for both languages followed. Participants who scored less than 80% were removed from the analysis.

	number	language D	gender D	gender TE	nº	example DP
1	SG	Dutch	M	M	2	**den** TOOTH
2	SG	Dutch	M	F	2	**den** DRESS
3	SG	Dutch	M	N	2	**den** CHALK
4	SG	Dutch	F	M	2	**de** BIKE
5	SG	Dutch	F	F	2	**de** DRAWER
6	SG	Dutch	F	N	2	**de** HOLE
7	SG	Dutch	N	M	2	**het** RENT
8	SG	Dutch	N	F	2	**het** ROOF
9	SG	Dutch	N	N	2	**het** SCARF
10	PL	Dutch			6	**de** DAYS
11	SG	English			6	THE **muts** 'hat'
12	PL	English			6	THE **landen** 'countries'

Table 5.2: Conditions for the Dutch-English gender stimuli

A screen with instructions in CS-mode was presented. The instructions were adapted from Koronkiewicz (2019a) and can be found in appendix a. The instructions were in code-switching mode to prime the participants for expecting the code-switching stimuli.

Stimuli were pseudo-randomly distributed over two blocks. Presentation of the blocks was randomised, as was presentation of the stimuli within the blocks. Participants were asked to judge a total of 72 sentences. These sentences were presented as Likert items. Participants could rate sentences between 1 ("completely unacceptable") and 7 ("completely acceptable"). The target stimuli represented half (36) of the total stimuli. Further included were 18 sentences used to investigate adverb placement and 18 control stimuli, developed according to the guidelines in Koronkiewicz (2019a). For details on the type of control stimuli, the

motivation for including them and analysis of the results, see appendix c.

After completing all judgments, participants were presented with a few post-test questions about their language use. They were asked to indicate their agreement with the following statements on a scale of 1–7. This measured a (self-reported) preference for one of the two languages (24a–b), attitudes towards CS (24c) and mixing behaviour (24c).

(24) a. In everyday conversation, I prefer speaking English.
b. In everyday conversation, I prefer speaking Dutch.
c. People should avoid mixing Dutch and English.
d. In everyday conversation, I mix Dutch and English.

They were also presented with monolingual control items to verify which gender they assigned to each of the nouns included in the set of stimuli, to verify that the variable under investigation was part of the participants' I-language, as recommended by Ebert and Koronkiewicz (2018). Participants that performed unexpectedly – i.e. giving the incorrect grammatical gender for the nouns used in the stimuli – were excluded from the analysis.

5.2.3 Results

Data was collected in early 2020 using Limesurvey software (LimeSurvey Development Team 2012). A total of 309 participants took the survey, 151 of whom completed it. Note that the high proportion of incomplete questionnaires is due to the automatic screening of participants. Participants who did not fit the profile were screened out. Participants who completed the survey took an average of 29 minutes to do so.

A number of participants were removed manually. Two participants didn't have at least one parent who was French-speaking or English-speaking and were removed. Five participants who reported only basic proficiency in English were removed. Partici-

pants reporting the use of Dutch (4) and English (5) on a monthly (or less than monthly) basis were excluded. Participants who scored under 80% proficiency for Dutch (4) and English (17) were also removed. Five participants were excluded because they rated *all* stimuli a 1. At this point 109 participants remained, 20 of whom were early bilinguals. My cut-off point for the early bilinguals was an age of aquisition (AoA) of both English and Dutch below 6, as it has been shown that such bilinguals give homogeneous acceptability judgments where CS data are concerned (López 2014). There were 49 late bilinguals, who learnt English after the age of 12. The participants that had an AoA of English between 6 and 12 (40) were eliminated, for maximal contrast between the two groups.

All results were analysed using R (R Core Team 2020). The data was prepared for analysis with the **tidyr** (Wickham and Henry 2020) and **dplyr** packages (Wickham et al. 2020). The visualisations were created with the **ggplot2** (Wickham 2016), **gridExtra** (Auguie 2017) and **likert** (Bryer and Speerschneider 2016) packages. Effect sizes were calculated using the **effectsize** package (Ben-Shachar et al. 2020).

Figure 5.2 presents an overview of the results, and these are summarised in table 5.3. This table shows the mean, standard deviation (StD), percentage of low (1–3), neutral (4) and high (5–7) ratings per condition. The first thing of note here is that the conditions with a neuter Dutch determiner (conditions 7, 8 and 9) are ranked rather low, contrary to the expectations from the literature, though condition 9, in which the TE of the noun was neuter, fares a bit better than the other two. In the top half of the table, we can find the conditions with a Dutch plural determiner (condition 10), an English determiner (11 and 12) and a feminine Dutch determiner (conditions 4, 5 and 6).

As discussed in section 3.2, all results will be analysed with both a parametric (t-test) and non-parametric (χ^2-test) tests. I'll also use statistical modelling to analyse the results.

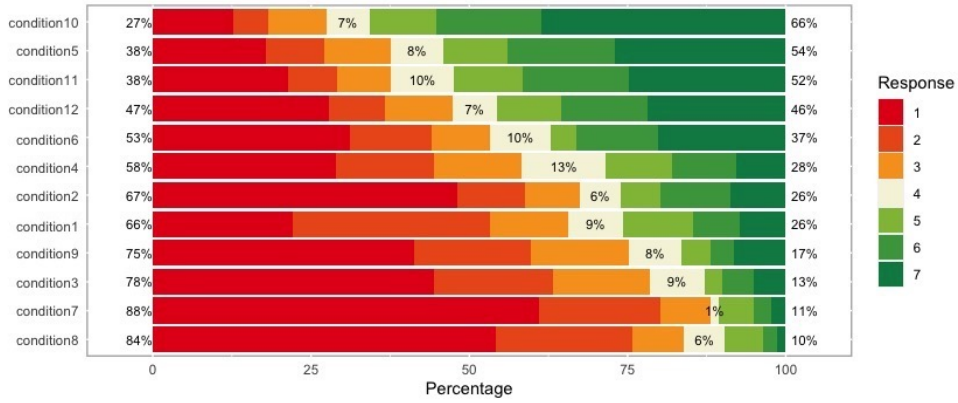


Figure 5.2: Overview of the ratings of the gender stimuli per condition for Dutch-English cs

condition	mean	StD	low	neutral	high
10	5.01	2.16	27.37	6.88	65.75
5	4.43	2.26	37.61	8.26	54.13
11	4.31	2.29	37.61	9.94	52.45
12	3.91	2.36	47.40	6.88	45.72
6	3.62	2.38	53.21	9.63	37.16
4	3.22	1.99	58.26	13.30	28.44
2	2.80	2.17	67.43	6.42	26.15
1	3.07	1.89	65.60	8.72	25.69
9	2.61	1.90	75.23	8.26	16.51
3	2.42	1.76	78.44	8.72	12.84
7	1.89	1.50	88.07	1.38	10.55
8	2.01	1.46	83.94	6.42	9.63

Table 5.3: Summary of the ratings of the gender stimuli per condition for Dutch-English cs

5.2.3.1 Tests of significance

I start with the predictions for the mixed DP asymmetry. In figure 5.3, we can see that there is only a minimal difference between the languages of the determiner, with the English determiner being slightly more preferred by both types of bilinguals. Note that this is the opposite of what is usually observed for the mixed-DP asymmetry. Usually, it is the determiner of the language *with* grammatical gender that is expected to be more acceptable.

Once we split up the results according to number of the embedded noun, a more telling picture emerges; both types of bilinguals have a preference for an English determiner in the singular, but for a Dutch determiner in the plural (figure 5.4). For both early (t = -5.90, df = 239.55, p-value = < 0.001 $|\chi^2$ = 44.29, df = 6, p-value = < 0.001) and late (t = -6.74, df = 643.77, p-value = < 0.001 $|\chi^2$ = 59.03, df = 6, p-value = < 0.001) bilinguals this difference is significant. This effect is large for the early bilinguals (Cohen's d = -0.76, 95% CI [-1.02, -0.50] |0.30, 95% CI [0.19, 0.38]) and medium for the late bilinguals (Cohen's d = -0.53, 95% CI [-0.69, -0.37] |Cramér's V = 0.23, 95% CI [0.16, 0.28]).[13] These confidence intervals (CIs) are pretty large, but the difference in effect size holds up at either end of the interval.

Let's take a look at gender agreement strategies. In order to do so, we need those stimuli that were Dutch sentences with an English embedded singular noun, as the Dutch determiner has no gender in the plural. In figure 5.5, you can find the mean ratings according to language of the determiner and bilingual type. Contrast this with figure 5.6, where the mean ratings are grouped according to gender of the determiner and gender of the translational equivalent of the noun. The left-hand chart in this figure (5.6) represents the early bilinguals and the right-hand chart the late bilinguals.

13 For interpretation of Cohen's d, I used the **interpret_d** function in the **effectsize** package, and for Cramér's V I used Cohen (1988, p 222).

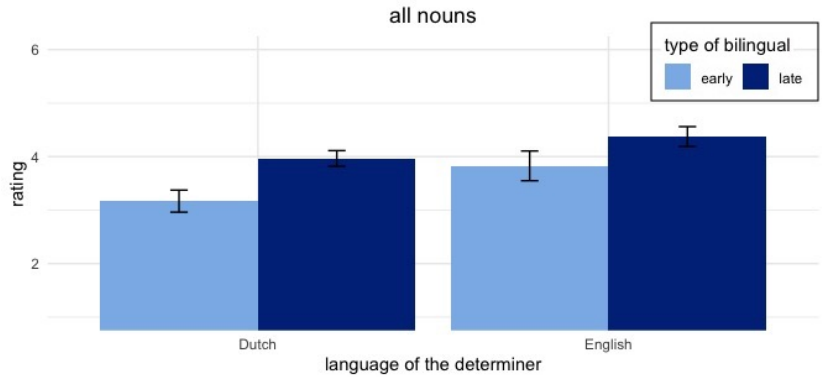


Figure 5.3: Dutch-English CS: mixed DP-asymmetry. Error bars represent the 95% CI

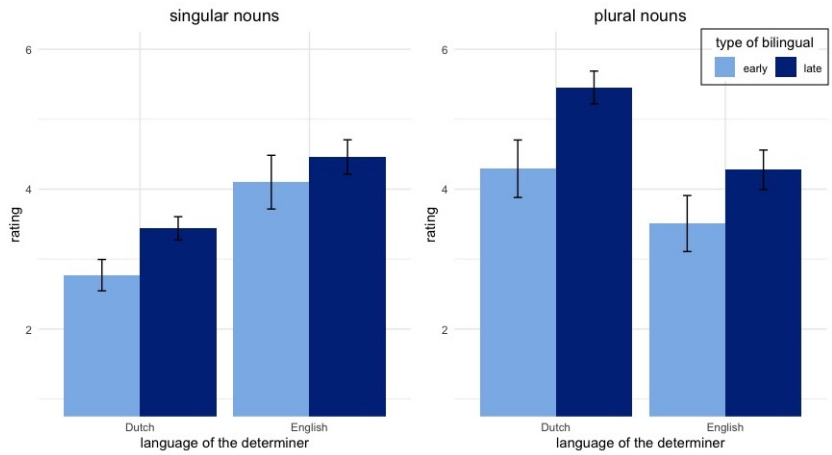


Figure 5.4: Dutch-English CS: singular vs plural DPs. Error bars represent the 95% CI

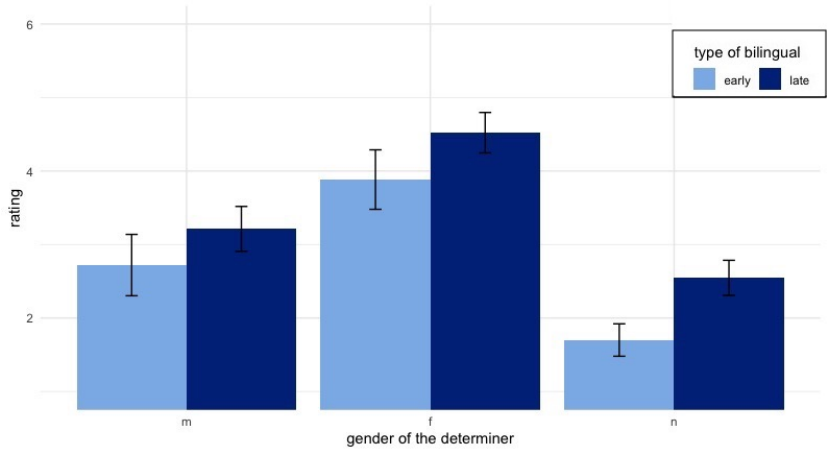


Figure 5.5: Dutch-English CS: determiner gender grouped by type of bilingual

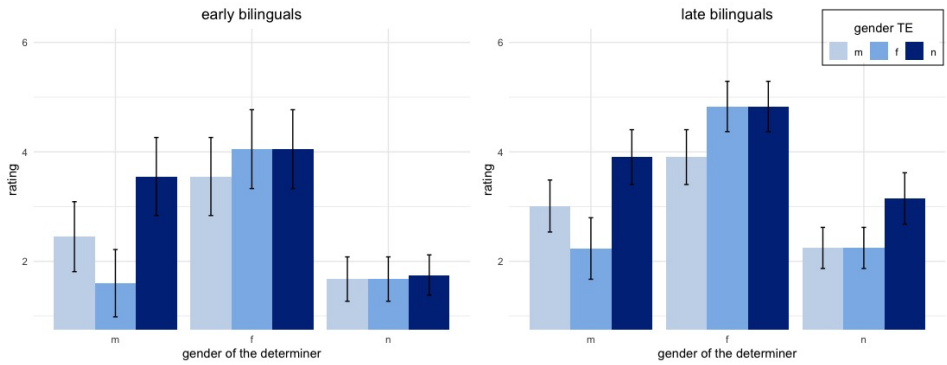


Figure 5.6: Dutch-English CS: determiner gender grouped by gender of the TE of the noun

	early bilinguals				late bilinguals			
	t-test		Wilcoxon-test		t-test		Wilcoxon-test	
	f	m	f	m	f	m	f	m
m	< 0.001	–	< 0.001	–	< 0.001	–	< 0.001	–
n	< 0.001	< 0.001	< 0.001	0.001	< 0.001	0.003	< 0.001	0.021

Table 5.4: p-values for the pairwise comparison tests

There are two surprising things in this figure. Firstly, unsurprising with regards to my judgments, but surprising from the point of view of the literature is that all bilinguals dislike the neuter determiner. In addition, contrary to both my and the literature's predictions the *feminine* determiner is the preferred one for both types of bilinguals. I will address this surprising result in the discussion section. All differences in rating between determiners were significant, as is shown in the results of the tests in table 5.4. Because multiple comparisons within the same dataset increase type I errors (Baayen 2008, p. 106), I used a pairwise t-test and Wilcoxon rank sum test with a Bonferroni-correction for multiple comparison.

The difference between the groups of bilinguals is not significant for the masculine (t = -1.89, df = 209.4, p-value = 0.06 |χ^2 = 5.60, df = 6, p-value = 0.47), but it is significant for the feminine (t = -2.58, df = 234.65, p-value = 0.01 |χ^2 = 12.94, df = 6, p-value = 0.044) and neuter (t = -5.14, df = 356.34, p-value = < 0.001 |χ^2 = 69.99, df = 6, p-value = < 0.001). For feminine this effect is small (Cohen's d = -0.34, 95% CI [-0.59, -0.08] |Cramérs V = 0.18, 95% CI [0.00, 0.25]), but for neuter it is medium (Cohen's d = -0.54, 95% CI [-0.76, -0.33] |Cramér's V = 0.38, 95% CI [0.28, 0.46]).

Looking at figure 5.6, we can see that the higher mean rating for neuter determiners in the late bilinguals is due to nouns with neuter as the gender of the TE, which the late bilinguals are sensitive to. This difference is statistically significant (t-test with

Bonferroni correction, p = 0.006 |Wilcoxon rank sum test, p = 0.01).

5.2.3.2 Modelling

The first research question I wanted to address concerned the mixed-DP asymmetry. In the maximal model I included a random slope and intercept for participant, but only a random intercept for item, as each participant only rated each item once (Judd et al. 2017, p 609). The fixed-effect structure of the maximal model included fixed effects for language of the determiner and number of the noun, type of bilingual, as well as the four extra-linguistic factors measured by the post-test questions (24). These extra-linguistic factors were included as we know from the literature that CS attitudes and behaviour may influence judgments of CS structures. A general preference for speaking one language over the other could be a possible influence in the choice of determiner, which is why it was included in the model.

I then used AIC_c and single term deletion to trim down the effect structure. None of the extra-linguistic factors were significant, and they did not improve the fit of the model, so they were eliminated. The two linguistic factors and the type of bilingual however, did improve the fit. I found that there was a triple interaction between the language of the determiner, number of the noun and the type of bilingual. The standard error (SE), p-values and effect sizes (standardised regression coefficient) with their CIs are shown for these factors, together with the significant interactions, in table 5.5.

As we can see, in general, when the noun is singular, the sentences were rated lower. This makes sense, as in the plural, there is no possibility for a conflict in gender agreement. We can see that, though by itself the language of the determiner is not a significant effect, this effect does interact with the number of the noun. This interaction was further analysed through the **emmeans** function of the **emmeans** package (Lenth 2020). I used

	SE	p-value	coefficient	95% CI
language D: English	0.789	0.113	-1.25	[-2.80, 0.30]
number N: singular	0.606	0.003	-1.81	[-3.00, -0.63]
late bilinguals	0.487	0.001	1.68	[0.73, 2.63]
language D*number N	0.933	0.006	2.57	[0.74, 4.40]
number N*late bilinguals	0.2489	0.001418	-0.79	[-1.28, -0.31]

Table 5.5: Ordinal model: Mixed-DP asymmetry in Dutch-English CS

the Bonferroni correction for multiple comparisons. In the plural, the Dutch determiner is preferred (1.45, SE = 0.716, p = 0.043), while in the singular, the English one is preferred (1.27, SE = 0.563, p = 0.025). This is not surprising considering figure 5.4.

Additionally, it seems like the late bilinguals tend to give a higher rating to the stimuli, something that was not necessarily predicted, but that I personally find unsurprising. This factor also interacted with number of the noun. In the plural in particular, the late bilinguals rated the items higher (1.54, SE = 0.562, p = 0.018).

Next, we turn to the gender agreement strategies. In order to do so, we only look at the data of singular embedded nouns preceded by a Dutch determiner, as Dutch plural determiners (and English ones) do not display gender agreement. For this subset of the data, I constructed a maximal model with a random slope and intercept for participant, but only a random intercept for item. The fixed-effect structure of the maximal model included fixed effects for gender of the determiner, whether the determiner was congruent with the translational equivalent of the noun, type of bilingual, as well as the four extra-linguistic factors mentioned above.

I then used AIC_c and single term deletion to trim down the effect structure. Of the extra-linguistic factors, only attitude towards CS (agreement with statement (24c)) improved the fit of the model. Whether the gender of the determiner was congruent

	SE	p-value	coefficient	95% CI
gender D: masculine	0.572	0.008	-1.51	[-2.63, -0.39]
gender D: neuter	0.600	< 0.001	-2.95	[-4.12, -1.77]
late bilinguals	0.54	0.021	1.24	[0.18, 2.30]
attitude towards CS	0.09	0.028	-0.44	[-0.83, -0.05]

Table 5.6: Ordinal model: Gender agreement strategies in Dutch-English CS

with the gender of the translational equivalent of the noun did not improve model fit, but the gender of the determiner did. Type of bilingual also improved the fit of the model. There were no interactions between these factors. Table 5.6 shows the standard error, p-value and effect size with its CIs for the factors of the final model.

From what we know of the literature on gender agreement in CS, it is quite surprising that the profile of the bilingual does not seem to matter, as neither group takes into account the gender of the TE of the noun. All bilinguals seem to use a default agreement strategy. The feminine seems to be the default gender: both the masculine and the neuter determiner had lower ratings, but the effect was almost twice as large for the neuter determiner. Contrary to the last model, there is an effect of attitude towards CS: if participants agreed with statement (24c), they rated sentences lower. However, as we can see, this effect is rather small. A possible reason for why these attitudes influence the ratings is that in all the items here, there was a possibility for grammatical conflict (the correct or incorrect determiner), whereas not all of the items tested with the previous model included such a conflict, which may account for the absence of an effect of CS-attitudes.

5.2.4 Discussion

In general, the results of this study were quite surprising and interesting. Overall, there was very little difference between the early and late bilinguals. Let's have a recap of what we might have expected based on the previous literature. As regards the mixed-DP asymmetry, we might expect a preference for the Dutch determiner as this is the determiner with grammatical features. As for gender agreement, we would expect an analogical gender strategy for the late bilinguals and a default strategy for the early bilinguals. And finally, regarding default gender, neuter seems to be the leading candidate based on the literature, but based on my own intuitions and informal observations masculine seemed more likely.

Let's start with the mixed-DP asymmetry. The data here do not correspond to the pattern which is often observed in bilingual corpus studies, where the language *with* grammatical gender tends to provide the determiner. This is perhaps not surprising, as this asymmetry has not yet been observed in AJTs. The preference for the Dutch determiner in the plural (where neither Dutch, nor English determiner has grammatical gender) is possibly due to the profile of the bilinguals, who tended to be Dutch dominant (53 out of 69 currently resided in Belgium). In the singular, the preference for the English determiner could be a gender avoidance strategy. Such a strategy has been reported for adjectives (but not determiners!) in Spanish-English CS in Northern Belize (Balam and Parafita Couto 2019).

As for gender agreement, the statistical modelling showed that the gender of the translational equivalent is not a factor that predicts ratings in either bilingual group for this population. Both groups seem to prefer the default strategy. What is the default gender then? It seems quite clear that neuter is not, as the neuter is disliked strongly. The feminine determiner is the determiner with the highest overall ratings, regardless of the gender of the

TE of the noun. This is *quite* surprising. However, I suspect there is influence from Standard Dutch. In SD the default determiner is the common determiner. In certain contact varieties of Northern Dutch, the common determiner is even pushing out the neuter. A case in point is the hit song *Waar is de meisje?* by the Dutch language hiphop-collective De Hoop (Delbeke and Sarens 2011). Besides the titular **meisje**, all nouns in the song are preceded by a common determiner, regardless of the grammatical gender they have in SD/BD. A selection is shown in (25a). There are also a plethora of English borrowings/code-switches, a selection of which is shown in (25b).

And yes, this led to a lot of complaining about the dilapidation of Dutch …

(25) Waar is
where is

a. de meisje, de vee-boer, de
the[C] girl[N] the[C] cattle-farmer[C/M] the[C]
toon-vrouw
show-woman[C/F]

b. de cashflow/backstage/podcast/airplane
the[C]

To verify this, I looked at participants that consistently used **ne(n)** instead of **een** for masculine Dutch nouns in the monolingual control stimuli that were included at the end of the survey. We might presume that they are robust Tussentaal speakers and therefore less influenced by SD. Since there are just eight early and eighteen late bilinguals, I won't do any statistical analysis on them, but the mean ratings for these participants are given in figure 5.7. In figure 5.8, we can see a comparison with the mean ratings by the participants who did not use **ne(n)** consistently, or even at all (i. e. the SD speakers). The figures look quite similar, and crucially, the robust Tussentaal speakers don't seem to have increased tolerance for the masculine determiner.

So where does this leave us? What is the unmarked gender in Dutch? The AJT presented in this chapter provides strong evi-

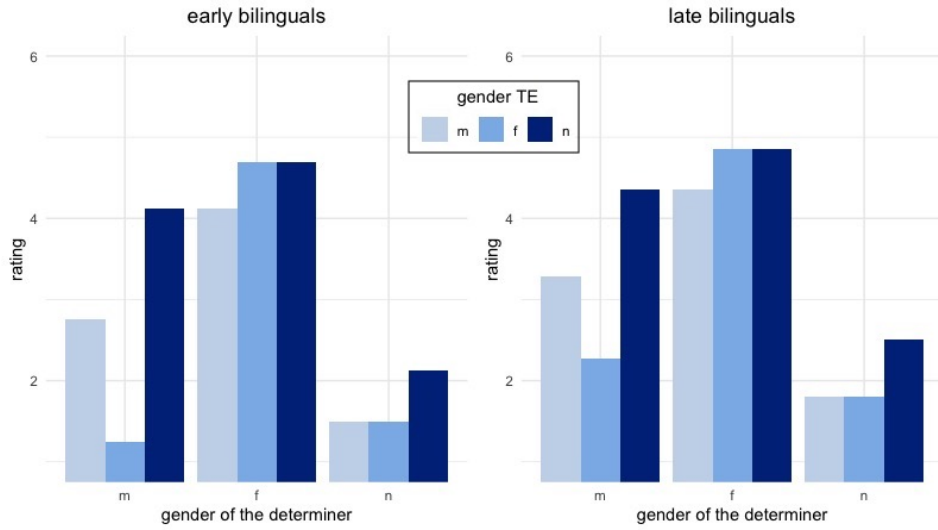


Figure 5.7: Dutch-English CS: determiner gender grouped by gender of the TE: Tussentaal speakers

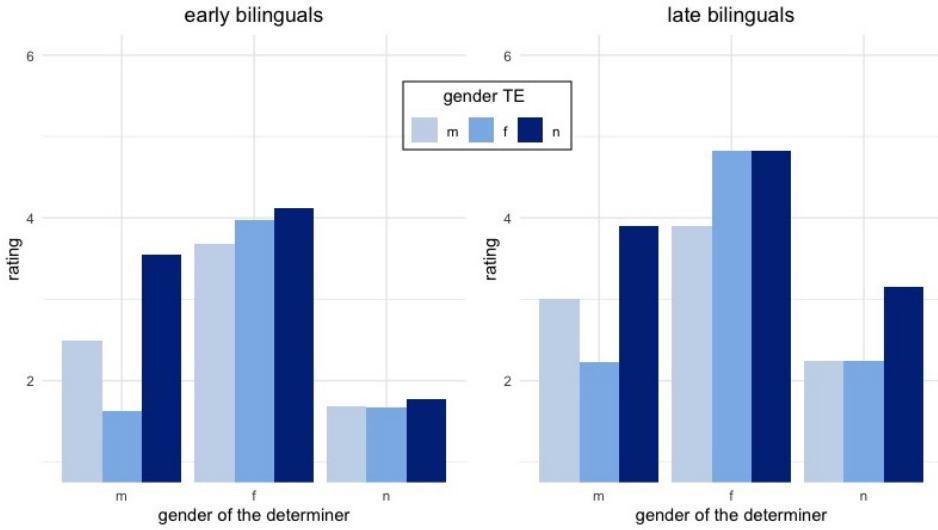


Figure 5.8: Dutch-English CS: determiner gender grouped by type of bilingual: SD speakers

	Belgian Dutch		Standard Dutch	
	gender	number	gender	number
den	∅	SG	—	—
de	-N	SG	∅	SG
het	+N	SG	+N	SG
de	∅	PL	∅	PL
het	∅	∅	∅	∅

Table 5.7: Feature inventory of definite determiners in Dutch

dence against neuter as the unmarked gender in Dutch. The fact that in SD the distinction between masculine and feminine has been neutralised, together with the disappearance of the neuter gender in some varieties of Northern Dutch lends further support that this neuter gender is marked. How can we account for neuter as an exceptional case default, as illustrated in examples (17)–(20)? The only way is assuming that the neuter that shows up in exceptional case defaults, and the neuter that shows up in the "regular" agreement system are not the same. In table 5.7 I show my proposed feature inventory (with regards to gender and number) for the definite determiners in BD and SD. As you can see, the system contains a fair amount of syncretism. The exceptional case default **het** is different from the neuter gender determiner **het** in that the former is unspecified for both gender and number, while the latter *is* marked for gender [+N] and number [+SG].

Possible additional evidence for the existence of two different neuter determiners in Dutch is found in the language acquisition literature. It seems that these two have different acquisitional paths. For children acquiring SD, Tsimpli and Hulk (2013) show that, with nominal goals, acquisition of the neuter determiner is delayed compared to the common determiner and that this common determiner is overgeneralised, indicating a default sta-

tus for the common determiner in SD (see also Roodenburg and Hulk 2009, among others). However, with non-nominal goals, the neuter determiner appears as early as the common determiner with nominal goals (Tsimpli and Hulk 2013, p 136).[14] The analysis of the exceptional case default as being unmarked for gender *and* number is how Picallo (2005, 2008) deals with the pronoun lo in Spanish, and this analysis is also adopted by Kramer (2015).

Based on the data I collected, I am unsure whether **de** or **den** is the unmarked determiner. It is impossible to tell whether what looks like a preference for the feminine determiner is actually that, or rather a preference for the SD common determiner. Based on my own intuitions,[15] and the informal data in (21) and figure 5.1, I am inclined to say that it is **den**. I cannot exclude however that it is **de**, or even that this may be different for different speakers. To make a definitive call on this issue, more research is required, with specific attention to avoiding a bias against Tussentaal.[16]

In summary, the results of the survey have shown no evidence for the mixed-DP asymmetry, a preference for the default gender agreement strategy for both early and late Dutch-English bilinguals, and have highlighted that the default gender in Belgian Dutch seems to be the feminine, though this may be due to influence from Standard Dutch, where the common gender is the default.

5.3 TESTING ENGLISH-FRENCH CS

Considering section 5.1, there are several questions we can ask when it comes to gender agreement to English nouns in English-French CS.

14 I would like to thank Miriam Greidanus Romaneli for pointing me to the relevant literature.

15 Additional consultation with BD native speakers confirm my personal intuitions.

16 One possible way of doing so may be to use aural, rather than written, stimuli.

1. Is there evidence for the mixed-DP asymmetry in English-French CS?
2. What gender agreement strategies exist?
3. Is there a difference in gender agreement strategies across groups of bilinguals?

To answer these questions, I developed an AJT, which is discussed in the following sections.

5.3.1 Materials

For this survey, 30 preterite – to avoid the simple present, as this tense has very different uses in French and English – ditransitive sentences with a single noun insertion were constructed. All inserted nouns were preceded by a definite determiner. This DP was the direct object, which occurred in sentence-final position. An effort was made to make sentences sound as natural as possible given the constraints.

Factors included in the stimulus design were:

- language of the embedded noun (which was the opposite of the rest of the clause)
- number of the embedded noun
- gender of the French determiner: M or F
- gender of the (TE of the) noun: M, F

This resulted in a total of 10 conditions, which are shown in table 5.8, each of which was presented in three lexicalisation, for a total of 30 sentences. For conditions 5–10, gender of the determiner was not a factor as neither the French plural determiner nor the English determiner have gender. In (26), some sample stimuli are presented. The full set of stimuli can be found in appendix b.

	number	language D	gender D	gender (TE) N	example DP
1	SG	French	M	M	*le* NECKLACE
2	SG	French	M	F	*le* DRESS
3	SG	French	F	M	*la* TREE
4	SG	French	F	F	*la* MOON
5	PL	French		M	*les* FEET
6	PL	French		F	*les* MISTAKES
7	SG	English		M	THE *vélo* 'bike'
8	SG	English		F	THE *boîte* 'box'
9	PL	English		M	THE *arbres* 'trees'
10	PL	English		F	THE *jambes* 'legs'

Table 5.8: Conditions for the English-French gender stimuli

(26) a. *Le client choisissait le* BAG.
the client chose the[M]
condition 1

b. *L'homme remplissait la* BOX.
the man filled the[F]
condition 4

5.3.2 Procedure

The procedure was the same as the one described in 5.2.2. However, in this survey, there was a total of 78 sentences, with target stimuli representing slightly over a third (30) of the total stimuli. The distractors consisted of 24 sentences used to test adverb position (see chapter 8 for details), and 24 control stimuli, developed according to the guidelines in Koronkiewicz (2019a).

5.3.3 Results

Data was collected in early 2020 using Limesurvey software. A total of 154 participants took the survey, 94 of whom completed it. Note that the high proportion of incomplete questionnaires is due to the automatic screening of participants. Participants who did not fit the profile were screened out. Participants who completed the survey took an average of 28.9 minutes.

A number of participants were removed manually. Two participants didn't have at least one parent who was French-speaking or English-speaking and were removed. Two participants who reported only basic proficiency in English were removed. Since all participants reported using both English and French on at least a weekly basis none were removed for infrequent use of either language. Participants who scored under 80% proficiency for French (1) and English (10) were also removed. One participant was excluded because they rated *all* stimuli a 1. At this point 78 participants remained, 24 of whom were early bilinguals (AoA

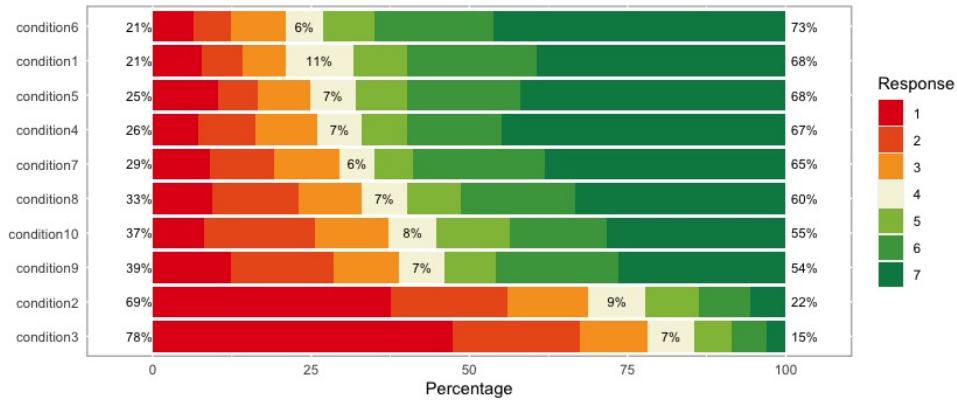


Figure 5.9: Overview of the ratings of the gender stimuli per condition for English-French CS

of both English and French below 6). For maximum contrast, I decided to compare these early bilinguals to those who acquired English after age 12. There were 19 participants in the latter group. I discarded the responses from participants who acquired English between the ages of 6 and 12 (35 participants). As each condition was presented 3 times, each condition received 72 judgments by early bilinguals and 57 judgments by late bilinguals.

Figure 5.9 presents an overview of the results, and these are summarised in table 5.9. This tables shows the mean, standard deviation (StD), percentage of low (1–3), neutral (4) and high (5–7) ratings per condition. At a first glance, it seems like the TE of the noun is an important factor: the two conditions for which the gender of the TE is incongruent with the gender of the determiner (conditions 2 and 3) were rated the lowest and rated markedly worse than the other conditions. These conditions either had a determiner without grammatical gender (conditions 5–10) or that had a determiner whose gender matched the gender of the TE of the noun.

condition	mean	StD	low	neutral	high
6	5.44	1.95	20.94	5.98	73.08
1	5.25	1.98	20.94	10.68	68.38
5	5.18	2.11	24.79	7.26	67.95
4	5.22	2.08	26.07	6.84	67.09
7	5.04	2.14	29.49	5.56	64.96
8	4.79	2.17	32.91	7.26	59.83
10	4.56	2.13	37.18	7.69	55.13
9	4.46	2.22	38.89	7.26	53.85
2	2.79	1.93	68.80	8.97	22.22
3	2.33	1.72	78.21	7.26	14.53

Table 5.9: Summary of the ratings of the gender stimuli per condition for English-French CS

5.3.3.1 Tests of significance

First, we will address the mixed-DP asymmetry. In figure 5.10, we can see that there is not really a preference for language of the determiner and that there is hardly any difference between the two types of bilinguals and that. Just like for the Dutch-English data, an interesting pattern is revealed by splitting up the nouns according to number. The rightmost graph in figure 5.11 shows the mean rating for singular nouns according to type of bilingual and language of the determiner. The leftmost graph in figure 5.11 shows the plural nouns. In the singular, there is a preference for the English determiner, while in the plural there is a preference for the French determiner, in particular with the late bilinguals.

The preference for the English determiner in the singular is statistically significant for both late (t = 2.2258, df = 285.36, p-value = 0.027 |χ^2 = 12.9, df = 6, p-value = 0.045) and early (t = -6.61, df = 384.78, p-value < 0.001 |χ^2 = 41.49, df = 6, p-value <

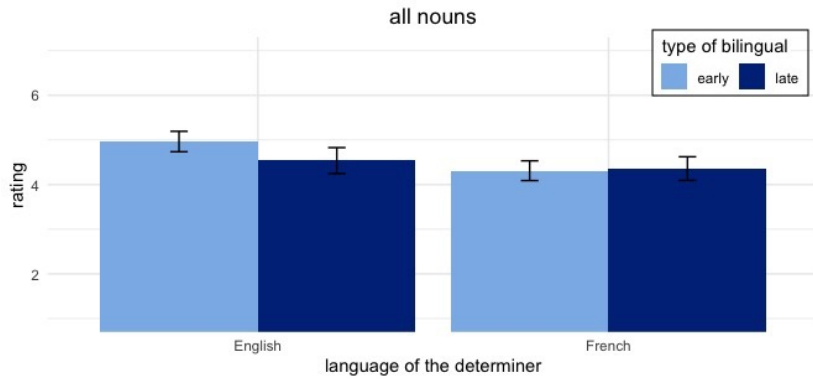


Figure 5.10: English-French CS: mixed DP-asymmetry

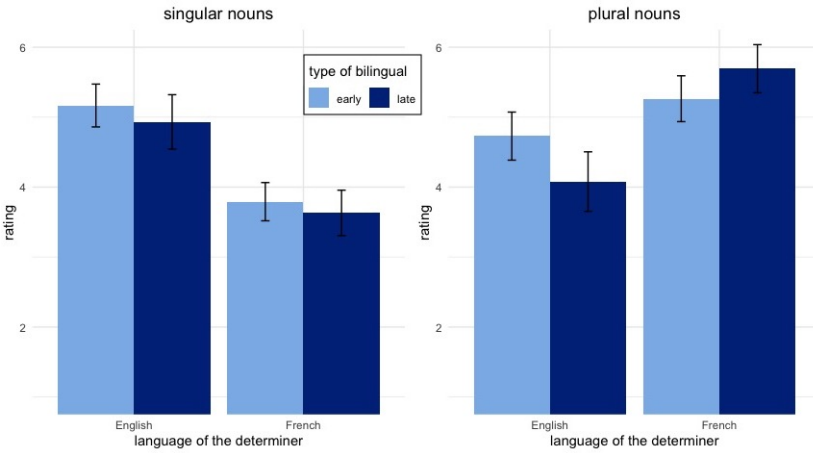


Figure 5.11: English-French CS: singular vs plural DPs

0.001) bilinguals, and the effects are large for the early bilinguals (Cohen's d = -0.67, 95% CI [-0.88, -0.47] |Cramér's V = 0.31, 95% [0.19, 0.39]) and small for the late bilinguals (Cohen's d = 0.26, 95% CI [0.03, 0.50]|Cramérs V = 0.21, 95% CI [0.00, 0.29]). The differences between the groups of bilinguals is not statistically significant (t = 1.22, df = 1177.4, p-value = 0.222 | χ^2 = 20.95, df = 6, p-value = 0.187).

In the plural, the preference for the French determiner is also statistically significant for both late (t = 5.85, df = 216.48, p-value

< 0.001 |χ^2 = 32.86, df = 6, p-value < 0.001) and early bilinguals (t = 2.23, df = 285.36, p-value = 0.027 |χ^2 = 12.896, df = 6, p-value = 0.045). The effect is small for the early bilinguals (Cohen's d = 0.26, 95% CI [0.03, 0.50] |Cramérs V = 0.21, 95% CI [0.00, 0.29]) and large for the late bilinguals Cohen's d = 0.79, 95% CI [0.52, 1.07] |Cramér's V = 0.38, 95% CI [0.21, 0.48]).

Let's move on to gender agreement. Just as in the Dutch-English data, this means we focus on the subset of data where the inserted noun is English singular. In figures 5.12 and 5.13 we can see the mean ratings according to language of the determiner (5.12) and with gender of the TE additionally factored in (5.13). We can see that the gender of the TE plays a role for both types of bilinguals: if the gender of the TE is masculine, the masculine determiner is preferred; if the gender of the TE is feminine, the feminine is preferred.

The early bilinguals seem to have a higher tolerance for the masculine determiner if the gender of the TE of the noun is feminine. This difference between bilingual types is significant (t = 2.01, df = 117.17, p-value = 0.047 |χ^2 = 13.05, df = 6, p-value = 0.042), but the effect seems to be small (Cohen's d = 0.37, 95% CI [0.01, 0.74] |Cramér's V = 0.32, 95% CI [0.00, 0.44]). I say "seems to be", as the CIs are even larger here than in the previous calculations. Within the early bilinguals, the difference in rating between the masculine determiner when the gender of the TE is feminine vs the feminine determiner when the gender of the TE is masculine is also statistically significant (t = 3.97, df = 127.33, < 0.001 |χ^2 = 23.97, df = 6, p-value < 0.001) and this is a large effect (Cohen's d = 0.70, 95% CI [0.34, 1.06] |Cramér's V = 0.37, 95% CI [0.17, 0.49]). For the late bilinguals this is not significant (t = 0.70, df = 111.32, p-value = 0.485 |χ^2 = 2.58, df = 6, p-value = 0.86), suggesting that there is a slight preference for default gender in the early bilinguals.

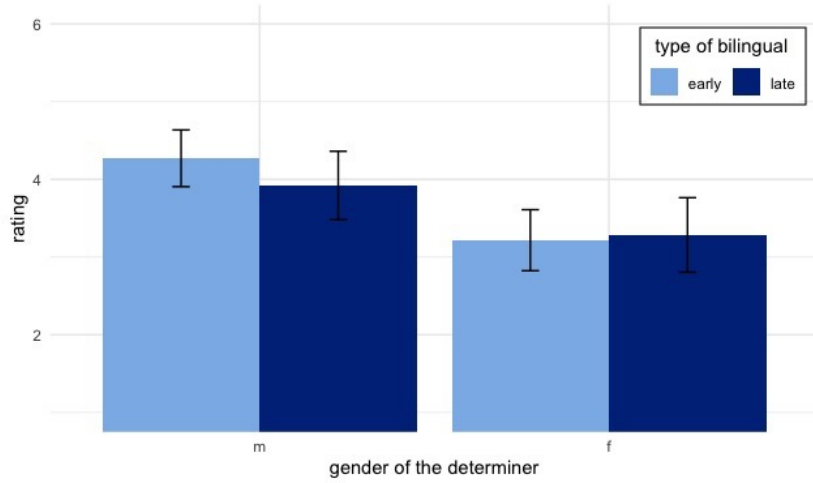


Figure 5.12: English-French CS: default gender grouped by type of bilingual

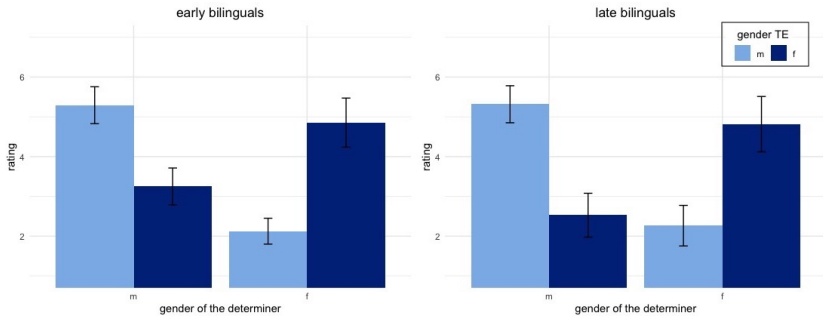


Figure 5.13: English-French CS: gender of the TE grouped by type of bilingual

5.3.3.2 Modelling

I first took a look at the mixed-DP asymmetry. In the maximal model I included a random slope and intercept for participant, but only a random intercept for item. The fixed-effect structure of the maximal model included fixed effects for language of the determiner and number of the noun, type of bilingual, as well

as the four extra-linguistic factors measured by the post-test questions (24).

I then used AIC_c and single term deletion to trim down the effect structure. None of the extra-linguistic factors were significant, though preference for speaking English did improve the fit of the model. Just like in the previous model, there was a triple interaction between the language of the determiner, number of the noun and the type of bilingual. The SEs, p-values and effect sizes with their CI, together with the significant interactions, are shown for these factors in table 5.10.

	SE	p-value	coefficient	95% CI
language D: French	0.779	0.441	0.60	[-0.93, 2.13]
number N: singular	0.683	0.338	0.65	[-0.69, 1.99]
late bilinguals	0.818	0.422	-0.66	[-2.26, 0.95]
preference for English	0.146	0.076	0.37	[-0.04, 0.79]
language D*number N	0.928	0.007	-2.52	[-4.34, -0.71]
language D*late bilinguals	0.6	0.008	1.58	[0.41, 2.76]
language D*number N*late bilinguals	0.459	< 0.001	-1.56	[-2.46, -0.66]

Table 5.10: Ordinal model: Mixed-DP asymmetry in English-French CS

None of the factors are significant by themselves, but the interactions between them are significant. Just like in the previous study, further analysis of the interaction between gender of the determiner and number of the noun showed that overall in the singular, the English determiner was preferred (1.88, SE = 0.605, p = 0.002), though in the plural, there was no significant preference (p = 0.121). As for the three-way interaction between language of the determiner, number of the noun and type of bilingual, further pairwise comparison (with Bonferroni correction) showed that the *late* bilinguals did tend to prefer the French determiners in the plural (2.179, SE = 0.812, p = 0.007), while the early bilinguals did not have such a preference (p = 0.441).

	SE	p-value	coefficient	95% CI
gender of the TE: congruent	0.5999	< 0.001	4.58	[3.60, 5.35]
preference for English	0.0070	< 0.001	0.41	[0.13, 0.94]

Table 5.11: Ordinal model: Gender agreement strategies in English-French CS

For the gender agreement strategies, I created a subset of the data that only included the sentences with French, singular determiners, as it is in these sentences that gender agreement is relevant. Again, the maximal model included a random slope and intercept for participant, but only a random intercept for item. The other factors that were included were gender of the determiner, congruence of the gender of the determiner with the gender of the TE and type of bilingual, as well as the four extra-linguistic factors measured by the post-test questions (24).

I then used the AIC_c and single term deletion to trim down the effect structure and the only factors that improved the fit of the model were congruence of the gender of the TE and a preference for speaking English. Both effects were significant, but the effect of preference for speaking English was quite small. The SE, p-values and effect sizes with their CI intervals are shown for these factors in table 5.11.

The results from this model indicate that all participants, regardless of their AoA of English prefer the analogical gender agreement strategy.

5.3.4 Discussion

With regards to the mixed-DP asymmetry, the results were very similar to the Dutch-English data. While overall there doesn't seem to be a preference for the French determiner over the English one (or vice versa), when the data is split according to number of the noun we can see a pattern: in the singular the

English determiner is preferred, again indicative of perhaps a gender avoidance strategy. In the plural there is no preference among the early bilinguals, but a preference for the French determiner in the late bilinguals.

As the late bilinguals are L_1 speakers of French and L_2 speakers of English, this translates into a preference for L_1 sentences with an embedded noun of their L_2, i. e. for sentences with a French matrix languages. The reason why this preference only shows up in the plural may be because the gender avoidance strategy is stronger than the preference for L_1 sentences with an embedded L_2 noun.

As for gender agreement, there was no difference in terms of bilingual profile: all bilinguals preferred sentences in which the gender of the determiner matched the gender of the TE of the noun.

5.4 CONCLUSION

In this chapter, I have investigated gender-agreement strategies in Dutch-English and English-French CS. Sentences with a determiner marked for grammatical gender had a lower mean rating than those that had a determiner not marked for grammatical gender for all types of bilinguals. This could be an indication of gender-avoidance. There was no evidence for the mixed-DP asymmetry, which is unsurprising, as it has not been observed in AJTs before.

As for gender agreement, the results were quite interesting. Recall that there is a robustly observed pattern in the literature that early bilinguals tend to go for the default gender strategy, whereas late bilinguals tend to assign code-switched nouns the gender of the translational equivalent. This pattern was observed for neither language pair under investigation in this chapter.

Overall, the Dutch-English bilinguals had a preference for default gender, regardless of their AoA of English. The default gender

seemed to be the feminine, and not neuter as predicted in the literature, nor masculine, which was the second likely candidate for default. This could be due to influence of Standard Dutch, where the common gender (whose determiners are formally identical to those of the feminine in BD) is the default. A possible reason why the late bilinguals don't seem to apply the analogical gender strategy may be because the gender system of Dutch is eroding, slowly but surely. If this is the case, it would be interesting to carry out a similar task with speakers of SD, where the erosion has progressed even further and speakers of German, where gender marking remains a lot richer. I suspect the gender of the TE may be more important for German-English bilinguals, whereas SD-English bilinguals may display an even stronger tendency towards the (common) default.

The English-French bilinguals on the other hand all had a preference for the analogical gender strategy. The early English-French bilinguals also showed sensitivity to the default gender agreement strategy, whereas the late English-French bilinguals did not. This could be seen as a tentative confirmation of the pattern documented in the literature.

From a methodological point of view, I have shown that the different statistical methods I applied led to drawing the same conclusions. The CIs of the effect sizes for the tests of significance were rather large in this case, especially in the case of the χ^2-tests. I did find the effect sizes to give an interesting indication and thought them a worthwhile addition to the reporting of statistical results. The statistical modelling gave insight into the data that the traditional tests did not, in particular when comparing the gender agreement strategies in the different populations. It also offers a more intuitive insight into complex phenomena such as gender agreement strategies, where there are different patterns that need to be compared, as they can all be modelled at once, whereas the traditional tests require subdivision of the data for each strategy one wants to compare.

6 LEVEL: EXPERT (ÜBERLINGUAL)

As has become apparent from the previous two chapters, the subject of grammatical gender is quite complex. When reading the literature on grammatical gender, one inevitably comes across authors who've also waded through the mire and have come out at the other end with some eloquent expressions of their ~~frustration~~ experience. Here is a collection of my favourites.

> In the domain of comparative grammar no subject is of greater interest than the origin of that mysterious grammatical mechanism known as grammatical gender.
>
> Flom (1903, p 1)

> La catégorie du genre pose un des problèmes les plus critiques de notre discipline.
>
> Hjelmslev (1956, p 213)

> Gender is the most puzzling of the grammatical categories. [...] it becomes more fascinating the more it is investigated.
>
> Corbett (1991, p 5)

> Indeed, gender is perhaps the only grammatical category that ever evoked passion.
>
> Matasović (2004, p 13)

> Linguistic Gender comes in many guises and serves many masters.
>
> Aikhenvald (2016, p 5)

Part III

WORD ORDER

Word order offers a striking illustration of how bilingual data provide a unique insight into linguistic structure. Contrasting the properties of languages involved in code-switching allows us to isolate which structural elements in particular influence word order.

7 VERB SECOND

7.1 INTRODUCTION

When it comes to using CS to investigate word order, Dutch and English form an interesting language pair, as they are minimally different from each other in two respects. In the first place, the underlying word order is different, with English being a subject verb object (SVO) language and Dutch a subject object verb (SOV) language (Koster 1975).[1] Secondly, while Dutch displays verb-second (V2) word order across the board in main clauses, in English V2 word order is limited to *wh*-questions and clauses with a fronted negation and negative polarity. I will set aside such "residual verb second" (Rizzi 1990) constructions for the remainder of this chapter. In (1) and (2), the relevant contrasts between Dutch and English canonical clauses is illustrated.

(1) **... dat Alex een appel at.**
... that Alex a appel eat[PST]
... THAT ALEX ATE AN APPLE.

(2) **Gisteren at Alex een appel.**
yesterday eat[PST] Alex a apple
YESTERDAY, ALEX ATE AN APPLE.

1 Dutch has also been analysed as an SVO language, most notably by Zwart (1997). This account, however "does not challenge Koster's conclusion that the main clause word order is derived from the embedded clause word order" (p 81), but argues that "Dutch and English differ with respect to object movement and verb movement." (p 86). This distinction is not relevant here, and I will assume Dutch is an SOV language.

The two differences just illustrated provide an ideal testing ground for investigating the role of C^0 – the functional head heading the clause – in determining word order.

In this chapter, I aim to show that current mainstream Generative analyses for the V2 phenomenon do not hold up to scrutiny when confronted with bilingual data.

7.2 VERB SECOND

In verb second languages such as Dutch (3), the finite verb moves to the second position of the clause. For Dutch, this movement can easily be discerned in sentences with a fronted adjunct (3b), or sentences with compound tense (3c), as Dutch – as mentioned above – is an SOV language. In English (4), the finite verb remains in post-subject position. Note that the V2 phenomenon is independent of underlying word order. Mainland Scandinavian languages for example are verb second SVO-languages (Holmberg 2015).

(3) Dutch: V2 – SOV word-order

a. Ik **zag** Kali (vandaag).
I see[PST] Kali

b. Vandaag **zag** ik Kali.
today see[PST] I Kali

c. Ik **heb** Kali gezien (vandaag).
I have Kali seen (today)

(4) English: no V2 – SVO word-order

a. I **saw** Kali.

b. Today, I **saw** Kali.

c. I **have** seen Kali.

In many V2 languages, the following asymmetry can be observed (5). In the main clause, the auxiliary is assumed to move to the

second position. In the subordinate clause, however, this auxiliary remains in its verb-final position.

(5) a. **Ik heb Kali gezien.**
I have Kali seen
b. **…dat ik Kali heb gezien.**
…that I Kali have seen
'(that) I have seen Kali.'

A host of different proposals exist to account for the asymmetry[2] shown in (5). For an overview of current approaches, I refer to Holmberg (2015). In this state of the art, Holmberg summarises the dominant Generative approach to V2 in the main clause as follows:

(6) a. a functional head in the left periphery (usually called C^0) attracts the finite verb
b. this functional head then attracts something (which may be the subject (3a), an adjunct (3b) or the direct object) to the specifier position of the complementiser phrase (CP)

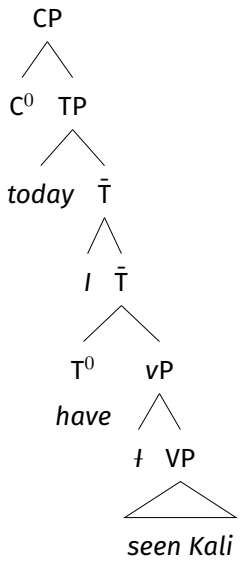
In the subordinate clause, the complementiser (e.g. **dat**) occupies the C^0 position and this prevents the finite verb from moving up, which accounts for the asymmetry shown in (5).

In non-V2 languages, the properties of the CP are such that in non-focus, declarative main clauses, C^0 does *not* attract anything, leaving the CP-layer empty. This is illustrated in (7). In Dutch (8), the finite verb has moved to C^0, prompting movement of the adverb *vandaag* to specCP. In English on the other hand, C^0 does not attract the finite verb and it remains in the head of the tense phrase (TP) , T^0.

2 Note that some languages do not display this asymmetry and have V2 word order in subordinate clauses as well, such as Yiddish and Icelandic. For an overview of analyses of such cases, see Holmberg (2015).

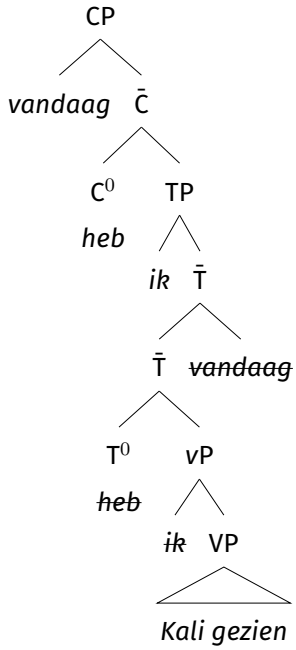
(7) a. [$_{\mathrm{CP}}$ [$_{\mathrm{TP}}$ Today [$_{\bar{\mathrm{T}}}$ I [$_{\bar{\mathrm{T}}}$ have]] [$_{\mathrm{vP}}$ ~~I~~ [$_{\mathrm{VP}}$ seen Kali]]]]

b.



(8) a. [$_{\mathrm{CP}}$ Vandaag [$_{\bar{\mathrm{C}}}$ heb] [$_{\mathrm{TP}}$ ik [$_{\bar{\mathrm{T}}}$ ~~heb~~] [$_{\bar{\mathrm{T}}}$ ~~vandaag~~] [$_{\mathrm{vP}}$ ~~ik~~ [$_{\mathrm{VP}}$ Kali gezien]]]]

b.



The summary in (6) characterises the common denominator of most approaches, but there are plenty of different analyses

available in the literature. Zwart (1993) is an example of an "asymmetric" analysis. Zwart argues that the verb moves to C *only* when a non-subject constituent is moved to specCP. Consequently, (4a) and (3a) look the same underlyingly in his analysis, with the verb in T^0 and the subject in specTP . Müller (2004) provides an account that differs even more substantially from the standard one. However, the trigger for V2 constructions is still an unlexicalised C^0, and Müller's account faces the same issues as the mainstream approach when it comes to CS data. Finally, there have also been implementations within a split CP analysis à la Rizzi (1997) (see Westergaard and Vangsnes 2005 for an example).

So while the exact details can differ from analysis to analysis they all share the same problem when it comes to accounting for bilingual data. The functional projection C^0 is *unlexicalised*, i. e. phonologically null. Consequently, the only way to tell what properties C^0 has is to look at the surface word order. If we have a surface V2 word order, we know that the C^0 has the properties of Dutch (8). If there is no V2 word order on the other hand, we are dealing with a C^0 with the properties of English (7).

As long as we are dealing with monolingual speakers, this remains unproblematic. However, as soon as we consider a (Dutch-English) bilingual speaker, a difficulty arises. Presumably, such a speaker would acquire both a C^0 of the "Dutch type" to derive Dutch sentences, and a C^0 of the "English type" to derive English sentences. Since C^0 is not phonologically realised by a word or morpheme in the declarative main clause, such a bilingual speaker could in principle use either of the C-heads when building their mixed-language structures. This would mean that if we follow this standard account for V2, no predictions are made in regards to preference between the word orders in (9). Depending on which C^0 is selected, either of these constructions can arise.[3]

3 As far as I can tell, there is nothing stopping a Dutch-English bilingual from using a "Dutch type" C^0 in constructions that have only English lexical items and the other way round. However, productions of this type are not attested, even in young children (Meisel 2004, p 100).

(9) a. TODAY, **ik zag Kali**.
I saw Kali

b. TODAY, **zag ik Kali**.
saw I Kali

This is exactly what is found by Jansen et al. (2012). From their analysis of corpora of 18 German-Romance bilingual children, the authors conclude that C^0 is indeed the determining factor for V2 word order (or lack thereof) in main clauses. They find that in mixed language constructions, both V2 and non-V2 word orders occur in the main clause (i.e. analogues to (9a) and (9b)). Note that the children they investigated varied significantly in age (from 1;9 to 5), exhibited varying levels of proficiency in both languages, and used their languages to different degrees.[4]

However, Jansen et al.'s findings do not correspond to my own intuitions as a sequential Dutch-English bilingual. An informal poll with simultaneous and sequential bilinguals in my immediate surroundings confirmed that there is a preference for (9b) over (9a).

Van Dulm (2007) finds a similar preference in her study on English-Afrikaans code-switching. In (10), an example of what she finds for focalisation (i.e. V2) constructions is shown. The V2 word order is preferred in sentences with an Afrikaans finite verb (10a). In (10b), the sentence has English (i.e. non-V2) word order, but has an Afrikaans finite verb, and is rejected by Van Dulm's participants.

(10) Afrikaans-ENGLISH CS Van Dulm (2007, p 60)

a. THE TALL PLASTIC CONTAINERS gebruik daardie kok
use that chef
FOR BROWN SUGAR.

4 The children had different "degrees" of bilingualism, as they were categorised anywhere between "balanced bilingual" and "extremely unbalanced bilinguals" – Jansen et al.'s terminology.

b. *THE TALL PLASTIC CONTAINERS daardie kok gebruik
that chef use
FOR BROWN SUGAR.

'The tall plastic containers, that chef uses for brown sugar.'

Van Dulm agrees with me that it seems that "the theory makes no prediction at all (or a prediction of 'anything goes') for the structure of code switched constructions" (p 50). I will return to her analysis in section 7.7.3

Outside of the syntactic literature, one study found a similar preference for analogues of (9b) over (9a) with Dutch-English (sequential) bilinguals. Kootstra et al. (2010) conducted a series of psycholinguistic experiments to investigate the influence of shared word order and alignment with a dialogue partner on word order choice in a language switching paradigm. The authors looked at word order in both main and subordinate clauses. For this language pair, in declarative main clauses with initial subject and no further adverbials, the *surface* word order is shared: SVO. In subordinate clauses and clauses with a fronted constituent, word order is not shared between English (SVO) and Dutch (SOV and V2, respectively). The authors found that – if possible – participants opted for constructions in which word order was shared, that is main clauses without fronted constituent. However, when they opted for a construction where shared word order was unavailable, V2 (main clauses) or SOV (subordinate clauses) was preferred when the finite verb was Dutch, but not when it was English. The data presented in their paper match my own intuitions for verb second, and provide interesting insights into word order in the subordinate clause as well, which will be discussed later in this chapter.

Kootstra et al. focus on language processing and are not concerned with accounting for their findings in a theoretical syntactic

framework. Their results seem to indicate that there *is* preference for word order (verb second or not) in mixed language constructions, something which is not predicted by the mainstream Generative point of view described in Holmberg (2015). There is, however, one paper that would predict their findings. Rambow and Santorini (1995) propose that in V2 languages, the verbal head itself causes the movement out of the TP , creating a head (C^0) with a Specifier position which must be filled by either the subject (as in (3a)) or (what they call) a topic (as in (3b)).[5] So rather than C^0 prompting V2 movement, it is V^0 that does. This proposal constitutes a radical departure from the mainstream accounts, and has been largely neglected in later treatments of V2. It is telling that Holmberg's (2015) state of the art makes no mention of Rambow and Santorini (1995).

In summary, while many different analyses have been proposed in the Generative literature for the verb second phenomenon, all but one share the idea that an *unlexicalised* functional head in the left periphery is responsible for V2 word order. A single paper, we have just seen, argues that it is the *verbal* head that causes this movement to a higher position. Bilingual data provide unique insights into this matter, which can broaden our understanding of the V2 phenomenon. In order to tease apart the different factors involved in V2 word order, I developed an experiment, which is described in section 7.5.

7.3 WORD ORDER IN THE SUBORDINATE CLAUSE

As argued in the previous section, the fact that C^0 is an unlexicalised element in the main clause leads to a lack of prediction for word-order preference in V2 constructions. On the mainstream Minimalist approach, depending on which C^0 (the English or Dutch one) is selected, either word order in (9) can surface.

5 For full details of the derivation, I refer to Rambow and Santorini (1995, p 280).

Yet, C^0 is not always unlexicalised. In subordinate clauses, the position is occupied by a complementiser. Hence, it might be interesting to see what determines word order (SVO vs SOV) in a mixed Dutch-English subordinate clause.

Contrary to what was observed for V2 constructions, there are multiple papers in the CS literature that address this issue. Two main points of view can be identified.[6] Either the language of T^0 determines word order, or the language of C^0 determines word order. Proponents of the former view (Chan 2008; MacSwan 2004; Mahootian 1993; Mahootian and Santorini 1996 among others) argue that any functional head determines the order of its complement. Since the verb phrase (VP) is the complement of T^0, as is illustrated in the simplified structure in (11), it is this head that determines the word order between verb and direct object. The findings in these studies dovetail with those in Van Dulm (2007) and Kootstra et al. (2010).

(11) $[_{\mathrm{CP}}\ C^0\ [_{\mathrm{TP}}$ subject $T^0\ [_{\mathrm{VP}}$ verb $[_{\mathrm{DP}}$ direct object $]]]]$

Contrastingly, Jansen et al. (2012) and Cantone (2007) contend that the complementiser (i.e. C^0) determines the grammatical properties – word order among these – of the subordinate clause. If they are on the right track, there is an expected preference for constructions such as (12a). If, on the other hand Chan (2008), MacSwan (2004), Mahootian (1993), and Mahootian and Santorini (1996) are correct, there is an expected preference for (12b).

(12) a. ALEX KNEW $[_{\mathrm{CP}}$ THAT $[_{\mathrm{TP}}$ **ik zag Kali**. $]]$
I saw Kali
embedded SVO order matches English C^0

6 This is not entirely the same debate as the one going on in the monolingual literature. Since a discussion of that debate point falls outside the scope of this chapter, I refer the reader to Chan (2008, pp 789-792) for a short overview of how word order between verb and object has been analysed since Kayne's antisymmetry hypothesis (Kayne 1994).

b. ALEX KNEW [$_{CP}$ THAT [$_{TP}$ **ik Kali zag**.]]
I Kali saw
embedded SOV order matches Dutch T^0

Whether T^0 or C^0 determines word order in the code-switched Dutch English clause was tested in an experiment which is described in section 7.6.

7.4 PUTTING THE THEORY TO THE TEST

The Dutch-English language pair provides the ideal testing ground to look at word order in both main and subordinate clauses. One of these languages is verb second, with underlying SOV word order. The other only displays residual V2 and is SVO.

The theories discussed in the previous sections were tested through an acceptability judgment task (AJT). We have seen that mainstream Generative accounts predict no preference for V2 vs non-V2 word order preference in Dutch-English code-switched clauses. We have also seen that one paper, Rambow and Santorini (1995), argues that it is the verbal head, rather than an unlexicalised C^0, that drives movement and results in V2 word order. Word order and language of the verb are then two factors that should influence acceptability judgments.[7] Whether or not language of the verb is correlated to word order will be tested in the experiment described in section 7.5.

For the subordinate clause, as we have seen, the main question is which functional head exactly determines the relative order of the verb and the object. In the literature discussed in section 7.3 two candidates can be discerned, C^0 and T^0. The experiment comparing these two predictions will be discussed in section 7.6.

7 Note that for the purposes of this chapter "language of T^0" and "language of the finite verb" are taken to be synonymous, as my AJT did not contain word-internal switching. For more about this, see section 7.7.1 and section 2.5.

7.5 MAIN CLAUSE

7.5.1 Procedure

To investigate what word order native bilinguals prefer, an acceptability judgment task was devised. The acceptability judgment task consisted in an online questionnaire developed using the survey software Qualtrics which participants took at home. The survey consisted of the following parts, in the following order:

- Start-up screen. Contained a welcome message displayed in both Dutch and English.
- Language background questionnaire (part I). Available in Dutch or English. Participants were able to choose in which language to complete the questionnaire.
- Written instructions for the acceptability judgment task, in code-switched text (for instructions, see appendix a).
- Three trial items.
- Four randomised blocks of 19 items, randomly presented.
- Stimuli were pseudo-randomly distributed over the 4 blocks.[8]
- Proficiency test. Ten items for Dutch and ten for English.
- Language background questionnaire (part II). Available in Dutch or English. Participants were able to choose what language to take the questionnaire in.

All stimuli were presented in their written form, rather than aurally. This written presentation avoided a potential pitfall of aural presentation, the fact that it is very easy to mistake the English complementiser (/ðæt/) with the Dutch one (/dɑt/). This distinction was crucial for the stimuli of the second experiment.

Respondents were also asked to provide linguistic background information. This background questionnaire was split over two parts, one presented before, and one presented after the test. The background question presented before the test contained

8 There was an equal representation of V2 and subordinate clause stimuli in each block. Conditions were evenly distributed across the blocks.

questions that allowed for screening out participants automatically. For example, if the language background questionnaire showed that the respondent no longer used both languages on a daily basis, their results were also excluded. The questions presented after the test asked questions that could possibly influence the ratings, such as their attitude towards code-switching (CS) (13b) and their own CS behaviour (13b).

(13) a. To what extent do you agree with the following statement: People should avoid mixing Dutch and English in the same conversation.
b. To what extent do you agree with the following statement: In everyday conversation, I keep the Dutch and English languages separate.

A small proficiency test was included. Results from respondents who scored below 80% on the proficiency test, and from bilinguals that acquired their L_2 after the age of six were excluded, as recommended by López (2014).

(14) a. People should avoid mixing Dutch and English.
b. In everyday conversation, I mix Dutch and English.

In the instruction block, respondents were asked to rate code-switched sentences, on a seven point scale, ranging from "completely unacceptable" to "completely acceptable".

7.5.2 Materials

For all stimuli, special care was taken to make sure that the switch point was unambiguous. Hence, no proper names occurred at the intended switch point. Switches between subject and verb were also avoided, as these are well known to be ungrammatical if the subject is a pronoun, as was first observed in Timm (1975). All stimuli were in Dutch-English CS mode.

All sentences were monotransitive and contained a fronted adjunct. The following factors were manipulated:

- word order: V2 or no V2
- language of the fronted constituent
- language of the verb
- language of the direct object

From the literature review in section 7.2, we know that language of the finite verb and word order can influence judgments. In addition, in order to create a variety of switch points, the language(s) of the direct object and fronted constituent were also manipulated. Combination of these factors resulted in 12 conditions (2*2*2*2 = 16 - 4 monolingual conditions). Table 7.1 shows an overview of these conditions with an example for each of them.

Each condition was presented in three lexicalisations to improve statistical power, yielding 36 test items, all with different lexicalisations. The 40 stimuli described in section 7.6.2 were used as the filler items. The full set of stimuli can be found in appendix b.

7.5.3 Predictions

The conditions in table 7.1 are ordered according to expected acceptability, with conditions likely to be rated highest at the top of the table. As the language of the object is not expected to influence word order (Koronkiewicz 2019a), conditions 6 and 7 are predicted to be judged acceptable. These sentences have only a code-switched object DP (**een schets** 'a sketch'; AN ASIAN DISH), and word order matches the language of the rest of the sentence.

Conditions 5 and 8 represent the mirror image of 6 and 7; except for the object DP (**hun huiswerk** 'their homework'; HER BIRTHDAY), they contain lexical material entirely from one lan-

		language of the			
	word order	fronted constituent	finite verb	direct object	example
6	no V2	English	English	Dutch	In the book Thelma draws **een schets**. a sketch
7	V2	Dutch	Dutch	English	**Wekelijks koken mijn ouders** an Asian dish. every week cook my parents
2	no V2	Dutch	English	Dutch	**Gisteren** the boy ate **een rijsttaart**. yesterday a rice cake
3	V2	English	Dutch	Dutch	Next year **bezoeken wij het Rijksmuseum**. visit we the "Rijksmuseum"
10	no V2	Dutch	English	English	**Op zondag** our son plays football. Sundays
11	V2	English	Dutch	English	In the garden **vangt de kat** a mouse. catches the cat
1	V2	Dutch	English	Dutch	**In die winkel** buys the dancer **haar jurken** in that shop her dresses
4	no V2	English	Dutch	Dutch	In the evening **hij kuste zijn dochter**. he kissed his daughter
9	V2	Dutch	English	English	**Op de trein** saw they a dog. on the train
12	no V2	English	Dutch	English	Every morning **de priester zegent** the holy water. the priest blesses
5	V2	English	English	Dutch	Sometimes forget students **hun huiswerk**. their homework
8	no V2	Dutch	Dutch	English	**In maart Janne viert** her birthday. in March Janne celebrates

Table 7.1: Examples sentences per condition: V2

guage, but word order as in the other. Consequently, they are expected to be judged unacceptable.

The interesting cases are represented by the conditions between the two dashed lines. The ones above the dotted line represent conditions in which the language of the finite verb and word order match, as in (15).

(15) CONDITION 2: **Gisteren** THE BOY ATE **een** **rijsttaart**.
yesterday a rice cake

The account in Rambow and Santorini (1995) predicts that these types of sentences should be preferred over the ones where the language of the finite verb (and consequently the verbal head) and word order do not match, as in (16).

(16) CONDITION 1: **In die winkel** BUYS THE DANCER **haar jurken**.
in that shop her dresses

However, if the mainstream account is correct, no such preference in acceptability judgments should arise.

7.5.4 Results

Data collection took place in the spring of 2017. A total of 100 responses were collected. Of those, 52 were discarded due to failure to complete the survey. Of the remaining 48, a further 34 were excluded, as they were responses by sequential – not simultaneous – bilinguals. Results from the remaining 14 participants were analysed. Three of these participants were male, eight female and the remaining three declined to answer the question about gender.

The results were analysed using R (R Core Team 2020). The data was prepared for analysis with the **tidyr** (Wickham and Henry 2020) and **dplyr** (Wickham et al. 2020) packages. The visualisations were created with the **likert** package (Bryer and

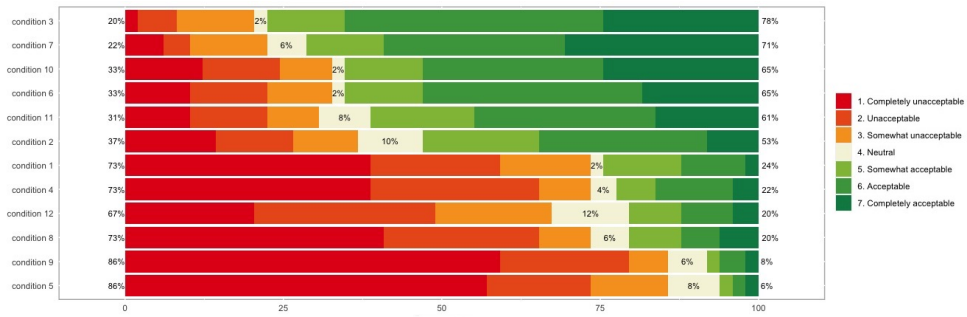


Figure 7.1: Overview of the ratings of the verb second stimuli per condition

Speerschneider 2016). Models were created with the **clmm** function in the **ordinal** package (Christensen 2019). Effect sizes were derived using the **effectsize package** (Ben-Shachar et al. 2020).

Each condition had three lexicalisations, so a total of 42 (3*14) judgments per condition were collected for this stimulus type. As discussed in section 3.2 I used both parametric and non-parametric tests to analyse the data. Figure 7.1 presents an overview of the results, and these are summarised in table 7.2. This tables shows the mean, standard deviation (StD), percentage of low (1–3), neutral (4) and high (5–7) ratings per condition.

In figure 7.1 and table 7.2, we can see that conditions in which word order matches the language of T^0 (conditions above the double line in table 7.1; 3, 7, 10, 6, 11, 2) are given a better rating than the ones where there is a mismatch between word order and T^0 (conditions above the double line in table 7.1; 1, 4, 12, 8, 9, 5).

7.5.4.1 Tests of significance

The difference between the conditions in which word order matches the language of T^0 and the ones where there is a mismatch between word order and T^0, which we observed in table 7.2 is not just impressionistic. If we compare the groupings of these categories, we find that they differ significantly (t = 13.34, df = 479.31,

condition	mean	StD	low	neutral	high
3	5.37	1.64	20.41	2.04	77.55
7	5.22	1.85	22.45	6.12	71.43
6	4.71	2.06	32.65	2.04	65.31
10	4.73	2.17	32.65	2.04	65.31
11	4.59	2.00	30.61	8.16	61.22
2	4.18	1.98	36.73	10.20	53.06
1	2.67	1.86	73.47	2.04	24.49
4	2.65	1.95	73.47	4.08	22.45
8	2.59	1.92	73.47	6.12	20.41
12	3.00	1.73	67.35	12.24	20.41
9	1.92	1.51	85.71	6.12	8.16
5	1.96	1.44	85.71	8.16	6.12

Table 7.2: Summary of the ratings of the verb second stimuli per condition

p-value < 0.001 |χ^2 = 135.75, df = 6, p-value < 0.001). This effect is very large (Cohen's d = 1.22, 95% CI [1.02, 1.41] |Cramér's V = 0.52, 95% CI [0.42, 0.60]).[9]

To see whether sentences are rated better if the language of the fronted constituent matched the word order, sentences in which word order matched the language of the fronted constituent (conditions 1, 4, 6, 7, 9 and 12) were grouped together. Here, the two tests gave conflicting results. The non-parametric test does not detect a significant effect (χ^2 = 11.1214, df = 6, p = 0.085) but the parametric t-test does (t = -3.18, df = 494.27, p-value = 0.002). While the t-test *does* come up with a significant difference in means for these two groupings, looking at the means themselves shows that the effect is in the wrong direction! When there is a *match* between the language of the fronted constituent and verb, the mean is 3.249, while if there is a *non-match*, the mean is 3.887. As for the effect size, it is estimated to be small (Cohen's d = -0.29, 95% CI [-0.46, -0.11]). The lower end of the CI even puts it at very small. Compare that to the impact of the verb that was discussed above, where even the lower end of the CI indicated a very large effect.

There is no statistically significant difference between a grouping of categories 6 and 7 (above the top dashed line in table 7.1, expected to be grammatical) and a grouping of categories 2, 3, 10 and 11 (above the dotted line) (t = 0.06, df = 171.58, p-value = 0.954 |χ^2 = 7.92, df = 6, p = 0.244). The same goes for groupings of categories 5 and 8 (below the bottom dashed line) versus a grouping of categories 1, 4, 9 and 12 (below the dotted line, expected to be ungrammatical) (t = -1.08, df = 182.67, p-value = 0.283 |χ^2 = 9.56, df = 6, p = 0.14). A random grouping of categories (such as even versus odd or first six versus last six) also resulted in non-significant effects, with p-values of 0.185 and 0.548 re-

9 The interpretation of the effect sizes for Cohen's d were obtained through the **interpret_d** function in the **effectsize** package, while for Cramér's V Cohen (1988, p 222) was used.

	SE	p-value	coefficient	95% CI
match word order and verb	0.6419	< 0.001	3.22	[1.96, 4.48]
CS behaviour	0.1864	0.0691	-0.69	[-1.44, 0.05]

Table 7.3: Main clause: Fixed effects for the ordinal mixed model

spectively (χ^2-test). These results suggest that it is only when the language of T^0 corresponds with word order that sentences are rated better.

7.5.4.2 Modelling

I created an ordinal mixed model that predicted the ratings on the basis of the other variables. In the maximal model, I included a random slope and intercept for participant, but only a random intercept for item, as each participant only rated each item once (Judd et al. 2017, p 609). The fixed-effect structure of the maximal model included fixed effects for variables that expressed a match between word order and verb on the one hand, and fronted constituent and direct object on the other. I included two extra-linguistic factors: self-reported CS behaviour and attitudes (i.e. agreement with the statements in (14)), as we know from the literature that CS attitudes and behaviour may influence judgments of CS structures.

I then used the AIC_c to trim down the effect structure. The model with the best fit included only effects for match between word order and verb and CS behaviour. While the latter only approached significance, removal resulted in a ΔAIC_c of 0.913, so I kept it in the model. The standard error (SE), p-value, effect size (standardised regression coefficient) with its CI for each of these factors are shown in table 7.3.

As we can see, when the word order and language of the verb match, the rating improves. If participants reported that they kept their languages separate (i.e. they agreed with the statement

in (14b)), their ratings were lower, but this effect was small and non-significant.

7.6 SUBORDINATE CLAUSE

7.6.1 Procedure

In this experiment, word order of the subordinate clause was investigated through an acceptability judgment task. Unlike in the main clause, the C^0 of the subordinate clause can be lexicalised by an overt complementiser. The data were collected as part of the same survey as the data for experiment one, so the procedure was identical to the one described in section 7.5.1.

7.6.2 Materials

As described in section 7.5.2, special care was taken to make sure that the switch point was unambiguous. Again, no proper names occurred at the intended switch point and switches between subject and verb were avoided. All stimuli were in CS mode.

All sentences of this type contained a main clause after which a switch would be made into the other language. One of the manipulated conditions was location of the switch point. There is some disagreement in the literature whether or not a switch between the complementiser and its complement is allowed. The functional head constraint (Belazi et al. 1994) prohibits a switch from occurring between a functional head and its complement. As C^0 is a functional head, this constraint predicts that switching between a complementiser and the TP will not occur. While much counter-evidence to the functional head constraint has been put forward (see among others Mahootian and Santorini (1996) for switches between a variety of heads and their complement and Bentahila and Davies, 1983, p 310 for C^0/TP switches),

González-Vilbazo (2005) argues that switching between C^0 and TP specifically is illicit, at least for the Spanish-German code-switching community he investigated. Hoot (2011), on the other hand, argues that a switch may be licit or illicit in English-Spanish CS depending on the grammatical constructions (see also Ebert and Hoot 2018; Sande Pineiro 2018). Because of these variable findings in the literature, it is important that the location of the switch – before or after the complementiser – be taken into account in any experiment designed to test these constructions, as is described in section 7.6.2.

The following factors were manipulated when constructing the test stimuli:

- language of the embedded finite verb
- switch before or after the complementiser (which determines the language of the complementiser)
- word order: SOV or SVO

Combinations of these factors resulted in 8 conditions (2*2*2). These combinations are shown in table 7.4 with an example of each condition.

Each condition was presented in five lexicalisations, yielding 40 test items. The 36 sentences described in section 7.5.2 functioned as fillers. The full set of stimuli can be found in appendix b.

7.6.3 Predictions

The factors described in section 7.6.2 allow one to isolate the possible influences on grammaticality. According to the previous literature, these can be:

- the relation between language of the complementiser (i.e. C^0) and word order (Cantone 2007; Jansen et al. 2012)
- whether the switch occurs before or after the complementiser (Belazi et al. 1994; González-Vilbazo 2005)

		language of the			
	word order	embedded verb	complementiser	switch	example
B	SOV	Dutch	Dutch	before C	MY FATHER SAW **dat de bal de muur raakte**. that the ball touched the wall
E	SVO	English	English	before C	**Marie dacht** THAT CATS EAT FISH. Marie thought
C	SVO	Dutch	English	after C	FELIX HOPED THAT **zijn zus vond een oplossing**. his sister found a solution
H	SOV	English	Dutch	after C	**De boer verwachtte dat** THE COW MILK GAVE. the farmer expected that
D	SOV	Dutch	English	after C	THE LAWYER KNEW THAT **de verdachte het goud stal**. the suspect the gold stole
G	SVO	English	Dutch	after C	**De student wist dat** THE TEACHER GAVE THE ANSWER. The student knew that
A	SVO	Dutch	Dutch	before C	HE HOPED **dat ze wonnen de wedstrijd**. that they won the game
F	SOV	English	English	before C	**Alfred raadde** THAT THE BUTLER JAMES KILLED. Alfred guessed

Table 7.4: Examples sentences per condition: subordinate clause

- the relation between language of the finite verb (i.e. T^0) and word order (Chan 2008; Mahootian and Santorini 1996; Van Dulm 2007)

As was the case in table 7.1, the conditions in table 7.4 are grouped according to predicted acceptability. Conditions B and E (above the top dashed line) are expected to be the most acceptable. There is no switching between C^0 and its complement TP , and both the language of C^0 and the language of T^0 match the word order.

Again, the conditions below the bottom dashed line (conditions A and F) represent the mirror image of those above the dotted line and are expected to be ungrammatical. While the switch occurs *before* C^0, the word order matches neither the language of C^0 nor the language of T^0.

While conditions C, D, G and H are all expected to fare badly if switching between C^0 and its complement is ruled out, these are the conditions which will allow us to differentiate between

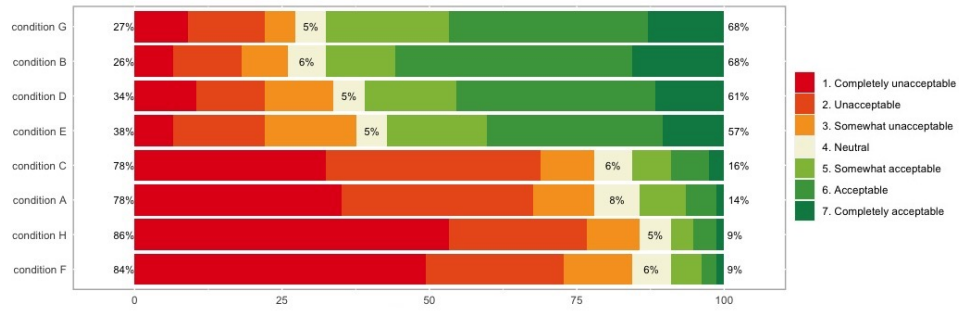


Figure 7.2: Overview of the ratings of the subordinate clause stimuli per condition

the influence of the complementiser and the finite verb. In conditions C and H, the word order matches the language of the complementiser, while in D and G it matches the language of the embedded finite verb. These two groups of conditions are separated by the dotted line. González-Vilbazo and López (2012) argue that there is a difference in acceptability between sentences separated by the dotted line, even though a C^0/TP switch is illicit.

7.6.4 Results

For information about number of responses, participant information and R packages see section 7.5.4. In this experiment, each condition had five lexicalisations, so a total of 70 (14*5) judgments per condition was collected for this stimulus type.

Figure 7.2 presents an overview of the results, and these are summarised in table 7.5. This tables shows the mean, StD, percentage of low (1–3), neutral (4) and high (5–7) ratings per condition.

Impressionistically, it seems again that sentences in which the language of T^0 matches the word order are rated better than those where there is a mismatch between T^0 and word order.

condition	mean	StD	low	neutral	high
B	4.88	1.88	25.97	6.49	67.53
G	4.69	1.92	27.27	5.19	67.53
D	4.52	1.96	33.77	5.19	61.04
E	4.42	1.88	37.66	5.19	57.14
C	2.48	1.66	77.92	6.49	15.58
A	2.42	1.58	77.92	7.79	14.29
F	2.08	1.45	84.42	6.49	9.09
H	2.00	1.47	85.71	5.19	9.09

Table 7.5: Summary of the ratings of the subordinate clause stimuli per condition

7.6.4.1 Tests of significance

A grouping of the categories G, B, D and E (word order and T^0 match) versus categories C, A, F and H (word order and T^0 mismatch), resulted in a significant difference in rating using both the (parametric) t-test and the (non-parametric) χ^2-test (t = 16.62, df = 506.3, p-value < 0.001 |χ^2 = 190.46, df = 6, p-value < 0.001). This effect was very large (Cohen's d = 1.48, 95% confidence interval (CI) [1.28, 1.67] |Cramér's V = 0.59, 95% CI [0.49, 0.66]).

In contrast, when categories were grouped according to whether or not C^0 matched the word order (categories B, C, E, and H vs A, D, F and G), the groups did not differ significantly on either test (t = -0.95, df = 548.9, p-value = 0.34 |χ^2 =3.09, df = 6, p-value = 0.8).

A grouping according to switch before (A, B, E and F) vs after (C, D, G and H) C^0 also did not result in a significant difference in rating (t = 0.14, df = 553.97, p-value = 0.893 |χ^2 = 1.74, df = 6, p-value = 0.942). It should be noted that the pronunciations of the Dutch (/dɑt/) and English (/ðæt/) complementisers are relatively similar, which may have played a role in the lack of support for the functional head constraint. However, since participants

	SE	p-value	coefficient	95% CI
match word order and verb	0.6258	< 0.001	3.74	[2.51, 4.97]
CS behaviour	0.1592	0.191	-0.42	[-1.06, 0.21]

Table 7.6: Subordinate clause: Fixed effects for the ordinal mixed model

were presented with written stimuli, I am assuming the role of phonology is limited, though it cannot be excluded.

7.6.4.2 Modelling

I created an ordinal mixed model that predicted the ratings on the basis of the other variables. In the maximal model, I included a random slope and intercept for participant, but only a random intercept for item. The fixed-effect structure of the maximal model included fixed effects for variables that expressed a match between word order and language of the verb on the one hand, and word order and language of the complementiser on the other. I included two extra-linguistic factors: self-reported CS behaviour and attitudes (i.e. agreement with the statements in (14)).

I then used the AIC_c to trim down the effect structure. The model with the best fit included fixed effects for match between the word order and language of the verb, and CS behaviour. While the latter was not significant, removal resulted in a ΔAIC_c of -0.472, so I kept it in the model. The SE, p-value, effect size with its CI are shown for each of these factors in table 7.6.

These results are very reminiscent of the results for the main clause. When the word order and language of the verb matched, the stimuli were rated higher. If participants reported that they kept their languages separate (i.e. they agreed with the statement in (14b)), their ratings were lower, but this effect was small and non-significant.

7.7 DISCUSSION & SUGGESTIONS FOR FURTHER RESEARCH

The results from the experiments described in this chapter are clear and robust; it is the language of the finite verb which determines the word order in both main and subordinate clauses in Dutch-English code-switching. This speaks in favour of accounts where the word order in subordinate clauses is determined by T^0 (Chan 2008; MacSwan 2004; Mahootian 1993; Mahootian and Santorini 1996), rather than C^0 (Cantone 2007; Jansen et al. 2012). In addition, since sentences with a switch before the complementiser were not rated significantly higher than sentences with a switch after, these results provide yet more counter-evidence to the functional head constraint (Belazi et al. 1994) and are consistent with the findings in Bentahila and Davies (1983).

For V2 clauses however, this is a surprising result. The mainstream Generative account predicts that either word order can be used, as was found in Jansen et al. (2012). The data used in that study – acquisition data – are of a very different type than the data used in this study – acceptability judgments by adult speakers. In addition, Jansen et al.'s data displayed great variability as the participants were a variety of ages, had variable proficiency in each of their languages, and were categorised as (being in the process of) achieving different degrees of bilingualism. This variability may go some way towards explaining our conflicting results. In Van Dulm (2007), data was collected through acceptability judgments, and results are consistent with the present study.

Incorporating the V2 results could lead us in two directions. One solution is to entirely re-think the usual way V2 is dealt with in the Generative literature. Simultaneous bilinguals show a significant preference for sentences that have a word order which is consistent with the language of the finite verb. One such analysis is provided by Rambow and Santorini (1995). On this proposal, it

is the verbal head that triggers V2 word order. However, Rambow and Santorini (1995) raises more questions than it answers. How or why exactly the verb moves in V2 languages, and not in non-V2 languages, is left unaddressed. Unfortunately, the paper has received no extensive follow-up to fine-tune the analysis.[10]

The other solution to account for these results is to rely on a mechanism which links (the properties of) T^0 and C^0. This could be accomplished in several ways. These possible solutions are examined in section 7.7.1 through 7.7.4

7.7.1 The PF interface condition

This constraint has been known as the Free Morpheme Constraint (Poplack 1980) and is currently formalised as the PF Interface Condition (PFIC) (MacSwan and Colina 2014) within Generative approaches to CS. The PFIC encapsulates the idea that the nature of phonological processing is such that switching from one phonological system to another word-internally is not possible. See section 2.5 for a more extensive explanation of the PFIC. In that section I discuss why I do not adopt the PFIC, but for the sake of argument, let's assume it does hold.

At a first glance, the the PFIC enables the Generative framework to make predictions about V2 word order. A finite verb consists of an affix (T^0) plus a verbal stem (V^0) which together form a so-called *complex head*. When this verb raises up further to C^0 (as in (8)), a new complex head (C^0+T^0+V^0) is created. Following the PFIC, this would mean that only a Dutch finite verb could raise to the left periphery in V2 constructions. If an English verb raised up to a Dutch C^0 (as it is only the Dutch C^0 that triggers V2 word

10 There is a brief follow-up by the first author in later joint work (Kinyon et al. 2006), in which the analysis in Rambow and Santorini (1995) is called a "pen-and-pencil analysis" (p 19) in which features "were introduced for descriptive reasons" (p 21). Unfortunately, most of the questions raised by the original paper do not receive a satisfying answer there.

order), a complex head with a word-internal switch would be created and the structure would be ill-formed.

I would like to argue that this is – in fact – not prevented from happening by the PFIC. Recall that this is a phonological constraint ("PF" stands for *Phonolical* Form, after all). The reason why word-internal switching is disallowed under the PFIC is that each lexical item is endowed with a phonological feature matrix. These matrices differ substantially across languages. Phonology is such that it applies to each word "in one fell swoop", which is not possible if a word contains elements whose phonologies are incompatible with one another (MacSwan and Colina 2014, 187ff).

However, it seems reasonable – even unavoidable – that a phonologically null functional projection has no phonology, and consequently, would never cause a violation of the PFIC. Assuming that an unlexicalised element has a phonology, despite being unpronounced, would be tantamount to introducing a "language label". True constraint-free approaches however, cannot rely on such language labels to explain code-switching patterns (MacSwan 2014).

7.7.2 The functional head constraint

I briefly mentioned the functional head constraint (Belazi et al. 1994) in section 7.6.2. Under this constraint, switching between a functional head and its complement is ruled out. This would mean that switching between C^0 and TP would be ruled out and would account for the results of the present study for the main clause. However, as mentioned previously, much evidence has been provided against the functional head constraint (for example Bentahila and Davies (1983) and Mahootian and Santorini (1996)). My results for the subordinate clause provide more counter-evidence to the functional head constraint, as condition G (a condition including a switch between C^0 and TP) was rated the highest of all conditions in that experiment.

In addition, the problem with this constraint in not solely empirical; it also fundamentally clashes with the spirit of the Minimalist Program (MP), as it is a CS-specific mechanism.

7.7.3 Feature Checking

As we saw in section 7.2, Van Dulm (2007) investigated word order in Afrikaans-English CS in structures involving verb raising. The relevant properties of Afrikaans parallel those of Dutch. These verb raising constructions she looked at include both main clause V2 constructions and embedded clauses, and her results are consistent with the ones in the present study; the finite verb determines word order.

Van Dulm's account is based on feature-checking. Recall from section 1.2 that syntactic features on functional heads drive movement: strong features need to be checked (hence, move) while weak features don't. Van Dulm attributes the fact that mainstream Generative approaches would seem to make "no prediction at all [...] for the structure of code switched constructions" (p 50) to a lack of specificity in the feature checking operation. If all lexical items can check the strong feature of the functional projection, anything goes. To solve this, she argues that the English and Afrikaans verb differ "in their ability to carry out particular checking operations" (Van Dulm 2007, p 51).

The word order differences between Afrikaans and English are derived as follows: "An Afrikaans finiteness phrase (FinP),[11] for example, may be proposed to possess a strong finiteness feature requiring checking, whereas an English FinP does not" (p 50). Crucially, this strong feature of the Afrikaans FinP can *only*

11 Van Dulm makes use of a split-CP structure. Under such an analysis, the CP is split into multiple functional projections. V2 word order is then derived as described in (6), but instead of a generic head in the left periphery attracting the verb, it is a specific head – such as FinP or focus phrase (FocP) – that does so, depending on the fronted element.

be checked by an Afrikaans verb. In other words, an English verb cannot raise into a V2 position, as it is unable to check the strong features in that position, resulting in the ungrammaticality of constructions like (10b), repeated here as (17).

(17) *THE TALL PLASTIC CONTAINERS daardie kok gebruik
that chef use
FOR BROWN SUGAR.

'The tall plastic containers, that chef uses for brown sugar.'

For Van Dulm, the fact that the functional head in the left periphery is not lexicalised does not pose a problem, as only verbs matching the properties of the unlexicalised head can raise to that position. In principle, this is a solution, but one that is not particularly attractive to me. Firstly, the constraints on the feature checking mechanism are not elaborated on and it is unclear what exactly Van Dulm understands by "the ability of the verb to check these features". Secondly, an approach like this assumes that verb second word order arises through the combination of two elements: a strong feature on C^0 and a verb that can check this feature. These two elements are not linked in Van Dulm's account; there is no reason for them to occur together, even though they always do in V2 languages. It seems that such an approach misses the aim of making use of "the minimal theoretical apparatus", as is the goal for researchers within the MP (Van Dulm 2007, p 26).

7.7.4 Feature Inheritance

Not to be confused with the neurobiological phenomenon of the same name

Another way of linking T^0 to C^0 is a mechanism called Feature Inheritance (FI). It was first proposed by Chomsky, who argued that since Tense and Agree features are inherent on C, while they

are derivative on T, "it makes sense to assume that Agree and Tense features [on T] are inherited from C" (Chomsky 2008, p 143).

There are several studies that rely on the FI mechanism to account for their data. Lochbihler and Mathieu (2016) and Jiménez-Fernández and Miyagawa (2014) argue that discourse features are inherited from C to T. Ulutas (2009) shows that subject case in Turkish embedded clauses also provides support for the FI mechanism. Several CS studies also provide empirical evidence. Shim (2021) relies on feature inheritance from *v* to ASP (though not from C to T) to for account word order phenomena in English-Japanese and English-Korean CS. López (2020) interprets the results in Sande Pineiro (2018) regarding pro-drop and Ebert and Hoot (2018) regarding the "that"-trace effect as providing support for FI from C to T.

However, the FI mechanism is not universally accepted in Generative grammar (López 2020, p 27). While Chomsky (2008) initially suggested complementiser agreement as a prime candidate for independent empirical confirmation of the FI mechanism (note 26), multiple authors have used data from complementiser agreement as evidence *against* FI – or rather, they argue that FI does not account for complementiser agreement patterns (see Diercks 2011 for Lubukusu (Bantu) and Haegeman and Van Koppen 2012 for West-Flemish (Germanic) language).

Additionally, Goto (2011) argues that FI is asymmetric and takes place in the embedded clause and not in the matrix clause, concluding *specifically* that "C-to-T Feature Inheritance does not take place in the V2 environment" (p 134). Indeed, the CS studies that provide evidence for C-to-T FI do rely on subordinate clause data, as does the one on Turkish (Ulutas 2009).

It seems that most studies that provide support for C-to-T FI rely on inheritance from a lexicalised C (or *v* in the case of Shim 2021). Hence, to me it remains unclear if C-to-T FI can be assumed to take place if C is unlexicalised. An alternative approach is developed by Gallego (2014). He argues that features are not in-

Surprise!!!

herited from C to T, but rather that "C and T are one and the very same unit in the lexicon" (p 41), which he calls the Feature Inheritance as Copying Thesis. I find this proposal more promising. It solves one of the issues of FI: the introduction of a new operation into the computational system. It also would work to account for the results of Ebert and Hoot (2018) and Sande Pineiro (2018). It would of course also explain my verb-second data, but the existence of CS between C and TP would seem to be problematic for Gallego's thesis. In addition, in that paper he restricts his attention to "situations in which both C and T are spelled-out" (p 67).

7.8 CONCLUSION

This chapter has provided evidence that in Dutch-English code-switched sentences, word order is aligned with the finite verb. For the subordinate clause, this is further evidence for accounts attributing verb-object order to T^0. This speaks in favour of accounts where the word order in subordinate clauses is determined by T^0 (Chan 2008; MacSwan 2004; Mahootian 1993; Mahootian and Santorini 1996), rather than C^0 (Cantone 2007; Jansen et al. 2012). For V2 constructions in the main clause, this seems to cast some doubt on the mainstream Generative account, which ascribes V2 word order to C^0. The V2 data in this chapter indicate that there is a close connection between T^0 and C^0. I have argued that neither the PFIC, nor Van Dulm's feature checking mechanism are attractive solutions to account for this close connection.

The best solution available in the literature is the mechanism of FI of Tense features from C^0 to T^0. However, I remain unconvinced that this mechanism straightforwardly applies in situations where C^0 is unlexicalised. A proposal along the lines of the one put forward by Gallego (2014), in which C^0 and T^0 are the same lexical unit, seems initially more promising, but it remains

an open question whether CS between C^0 and T^0 is compatible with that proposal. Hence, unless FI can be robustly shown to apply in main clauses (with an unlexicalised C^0), the data in this chapter present a challenge to Generative approaches to V2.

Additionally, I showed that, for my data at least, parametric and non-parametric statistical tests give you the same results. Only in one case did the parametric test detect an effect that was not found by the χ^2-test. Statistical modelling showed that the significant result of the t-test was in error.

8 ADVERB PLACEMENT

> Bovendien staat in bijna elke voorbeeldzin het adjectief ook op een verkeerde plaats, waardoor je als lezer geneigd bent om veel zinnen niet acceptabel te vinden.
>
> — participant 104

8.1 INTRODUCTION

When it comes to the position of the adverb in the clause, even closely related languages can behave differently. Take a look at the sentences in (1) and (2). In each of these sentences, an adverb intervenes between the verb and its direct object. In (1), the adverb is an adverb of frequency, but note that manner adverbs can also occur in this position. In (2) on the other hand, the adverb is an adverb of time. Note that the sentences in (1) and (2) aren't strict translational equivalents, but are chosen to be structurally parallel.

(1) a. **Ártemis leest vaak een boek**.
b. *Ártemis lit souvent un livre*.
c. *Ártemis reads often a book.

(2) a. **Ártemis las gisteren een boek**.
b. ?*Ártemis lisait hier un livre*.
c. *Ártemis read yesterday a book.

In English (the c. sentences), neither type of adverb can intervene between the verb and its direct object, while in Dutch (the a. sentences), both types of adverbs are perfectly possible in that position. Finally, French (the b. sentences) presents an in-

termediate case; manner adverbs are allowed (1b), while time adverbs result in a marked sentence (2b).[1]

Stowell (1981) is the first account of the observation in (1) and (2). He explains the difference between English and Romance languages in terms of case assignment: in order for the verb to assign case to its direct object, the verb and direct object need to be adjacent to each other. He dubs this requirement the Adjacency Condition (AC). In this chapter, I will use the term "AC violation" descriptively for structures in which the verb and its direct object are not adjacent to each other, even though an (apparent) violation of the AC does not result in ungrammaticality in all languages.

In English, there is a strict AC, which is why no adverbs are allowed to intervene between verb and direct object. In the Romance languages on the other hand, manner/frequency adverbs are supposedly "invisible for the purposes of Case assignment" (Stowell 1981, p 114), which is why an intervening manner/frequency adverb doesn't result in an ungrammatical sentence in these languages.[2] In Dutch then, he argues, no adverb causes an AC violation, as it is a V2 language. The direct object *is* in fact adjacent to the verb, albeit as a lower, unpronounced copy, and there is no problem with case assignment (p 117). This is illustrated in (3). For more information on V2 word order in Dutch, see section 7.2.

(3) **Ártemis las gisteren een boek ~~las~~**.
Ártemis read yesterday a book ~~read~~

1 I informally consulted ±25 native speakers, none of whom was willing to commit to the ungrammaticality of (2b). Most judged it as marked.

2 Stowell goes on to say that "[s]pecifically, we might hypothesize that Case assignment in Italian applies to an abstract representation of $\bar{X}$, where only the head and its arguments appear. [...] English and Italian would then differ according to whether Case assignment applies on the Argument-Projection" (p 114).

The adjacency-based account proposed by Stowell (1981) has largely been abandoned by accounts based on verb movement. For example, Pollock (1989), attributes these differences between French and English to differences in verb movement. He argues that English and French have the same underlying structure (4). The VP has an optional position for VP-internal adverbs.

(4) $[_{\text{IP}}$ NP I $[_{\text{VP}}$ (Adv) V ...]]

In French, the finite verb obligatorily raises to Infl (5), while, in English, this only occurs for auxiliaries (6).

(5) a. $[_{\text{IP}}$ *Ártemis est* $[_{\text{VP}}$ *souvent* ~~*est*~~ *heureuse*]] Aux-to-I
b. $[_{\text{IP}}$ *Ártemis lit* $[_{\text{VP}}$ *souvent* ~~*lit*~~ *un livre*]] V-to-I
c. *$[_{\text{IP}}$ *Ártemis* $[_{\text{VP}}$ *souvent lit un livre*]] *no V-raising

(6) a. $[_{\text{IP}}$ ÁRTEMIS IS $[_{\text{VP}}$ OFTEN ~~IS~~ HAPPY]] Aux-to-I
b. *$[_{\text{IP}}$ ÁRTEMIS READS $[_{\text{VP}}$ OFTEN ~~READS~~ A BOOK]] *V-to-I
c. $[_{\text{IP}}$ ÁRTEMIS $[_{\text{VP}}$ OFTEN READS A BOOK]] no V-raising

So what is at the root of the difference between French (2b) and Dutch (2a)? A difference in the level of attachment of the adverbs. The manner and frequency adverbs are VP-internal, while the others are VP-external.[3] In French, the VP-external adverbs occur to the right of the VP, while in Dutch, both VP-internal and external adverbs have their default position to the left of the OV complex (Koster 1975, p 123, Broekhuis 2013, p 476). This is shown in (7) and (8).

(7) VP internal adverbs

a. $[_{\text{IP}}$ NP I $[_{\text{VP}}$ (Adv) V O]] French, *souvent*
b. $[_{\text{CP}}$ NP C $[_{\text{IP}}$ I [(Adv) $[_{\text{VP}}$ O V]]] Dutch, **vaak**

3 This distinction between VP-internal and VP-external adverbs is also known as the distinction between VP adverbs and clausal/sentential adverbs (Laenzlinger 1998, p 42).

(8) VP external adverbs

a. [$_{IP}$ NP I [$_{VP}$ V O] (Adv)] French, *hier*

b. [$_{CP}$ NP C [$_{IP}$ I (Adv) [$_{VP}$ O V]]] Dutch, **gisteren**

In (9), an example of the abstract structure in (7) is given. This also shows the V-to-I movement for French and the V-to-I-to-C movement (V2 word order) for Dutch. In (10), we have an illustration of (8).

(9) a. [$_{IP}$ *Ártemis lit* [$_{VP}$ *souvent ~~lit~~ un livre*]]

b. [$_{CP}$ **Ártemis leest** [$_{IP}$ **~~leest~~** [$_{VP}$ **vaak een boek ~~leest~~**]]]

(10) a. [$_{IP}$ *Ártemis lisait* [$_{VP}$ *~~lisait~~ un livre*] *hier*]

b. [$_{CP}$ **Ártemis las** [$_{IP}$ **~~las~~ gisteren** [$_{VP}$ **een boek ~~las~~**]]]

We will set aside these VP-external adverbs (10) and focus on the VP-internal adverbs (9), as it it these that differ in distribution between French and Dutch on the one hand and English on the other.

8.2 DUTCH: TO ADVERB OR NOT TO ADVERB

I follow Broekhuis (2013) in the assumption that Dutch does not have a morphosyntactic category of adverb. The adverbial function is fulfilled by adjectives in Dutch.

(11) De auto rijdt snel.
'The car drives fast.'

You can easily show that words like **snel** are adjectives, as they can be used predicatively (12a) as well as attributively and they take adjectival agreement when in the latter position (12b).

(12) a. Dat was snel!
'That was fast!'

b. Dat is ne snell-en auto.
that is a[M] fast-M car[M]
'That's a fast car.'

There are a handful of cases for which their morphosyntactic status as adjective is less indisputable. For discussion, I refer the reader to Broekhuis (2013, p 464ff) and references therein, but the lexical items selected in the experiment below are all unambiguous adjectives used in so-called adverbial functions.

Since English and French *do* possess the morphosyntactic category of adverbs, I will also use "adverb" to refer to Dutch lexical items. Whenever this is done, it is to be understood as "adverbially used adjective".

8.3 THE ADVERB AND CODE-SWITCHING

There are to my knowledge only a few studies that specifically investigate adverb placement in CS. Stadthagen-González et al. (2018) (via both a classic AJT and two-alternative forced-choice task (2AFC)[4]) and Koronkiewicz (2019b) (via an AJT) investigated preferences for adverb placement in Spanish-English CS.[5]

Stadthagen-González et al. (2018) find that the acceptability of sentences containing AC violations was not predictable on the

4 A 2AFC consists in presenting participants with a pair of sentences and forcing them to choose the most acceptable one. Data from this task were analysed using a method derived from Thurstone's law of comparative judgment.

5 Note that Spanish differs from French, as it is not a pure verb-raising language, since both word orders in (i) are possible (Camacho and Sánchez 2017). The equivalent of (ia) does not freely occur in French.

(i) a. Juana siempre lee libros.
J. always reads books
b. Juana lee siempre libros.
J. reads always books

Camacho and Sánchez (2017, p 53)

basis of any one constituent, but rather that "all the constituents within the domain of the dependency contribute to the perceived (un)acceptability of the construction" (p 88). They found that constructions with an AC violation were improved by having a Spanish verb if either (or both) the direct object or adverb were in Spanish. However if neither of those constituents was in Spanish, a Spanish verb did not improve acceptability.

In a conference presentation, Koronkiewicz (2019b) reported on a study conducted on adverb position in Spanish-English CS. Whereas Stadthagen-González et al. (2018) tested only sentences *with* an AC violation, Koronkiewicz included other word orders as well. He found that the order S-ADV-V-O (12a) is always preferred. This makes sense, as this option is allowed in both English and Spanish. Whether or not the order S-V-ADV-O (12b) was judged acceptable was not predicted by the language of the verb, but rather, "if anything", by the language of the adverb.

Van Dulm (2007) (via AJT), on the other hand, looked at Afrikaans-English code-switching. The relevant properties of Afrikaans parallel those of Dutch discussed in section 8.1. The Afrikaans verb moves to a head in the left periphery and the adverb ends up between the verb and the direct object, as in (13), which parallels (1a).

(13) Hy koop dikwels lekkers.
he buy often sweet-PL
Afrikaans – Van Dulm (2007, p 54)

She found that structures containing an AC violation were only judged as acceptable if the verb was in the language that allows AC violations (Afrikaans). If the verb was English, structures without an AC violation were preferred.

These three studies present a puzzling picture, as they each yield different findings with regards to adverb position in CS. Van Dulm (2007) finds that the language of the verb predicts word order; Koronkiewicz (2019b) draws a tentative conclusion

that it is the language of the adverb; and Stadthagen-González et al. (2018) find that it is the language(s) of a combination of constituents that predict(s) acceptability of sentences with an AC violation.

In this chapter, I will add to the limited number of studies on this topic by investigating the AC violation in English-French and Dutch-English CS. The surveys I developed for that purpose and their results are discussed in the following sections.

8.4 TESTING DUTCH-ENGLISH

8.4.1 Stimulus-design

Nine frequency and manner adverbs were selected in Dutch and English.

- QUICKLY
- DAILY
- REGULARLY
- THOROUGHLY
- PATIENTLY
- OFTEN
- EASILY
- LOUDLY
- SOFTLY

- **snel**
- **dagelijks**
- **regelmatig**
- **grondig**
- **geduldig**
- **vaak**
- **gemakkelijk**
- **luid**
- **zacht**

I constructed 18 code-switched sentences. These 18 sentences consisted in 3 lexicalisations of 6 conditions. In all sentences, the adverb was placed between the verb and the direct object. The manipulated factors were:

- language of the adverb
- language of the direct object
- language of the verb

This 2*2*2 design yielded 8 conditions, two of whom were monolingual. These were excluded, so 6 conditions remained, which are presented together with an example in table 8.1. There was never a switch between subject and verb.

	language of the			
	verb	adverb	direct object	example
1	Dutch	English	Dutch	**Inge poetste** DAILY **haar tanden**. Inge brushed her teeth
2	Dutch	Dutch	English	**Ik verstopte snel** THE EASTER EGGS I hid quickly
3	Dutch	English	English	**De mannen bouwden** PATIENTLY A HOUSE OF CARDS. the men built
4	English	Dutch	Dutch	MY MOTHER EXAMINED **grondig de krant**. thoroughly the newspaper
5	English	English	Dutch	THE TEACHER SPOTTED EASILY **de fout**. the mistake
6	English	Dutch	English	MY HUSBAND FOUND **gemakkelijk** A PRESENT. easily

Table 8.1: Adverbs in Dutch-English CS: conditions

All sentences were simple transitive sentences in the preterite. They all had an animate subject and inanimate object. Close cognates were avoided, so that switch points were always immediately clear.

8.4.2 Procedure

The survey started with a short background questionnaire. At this early stage, some participants were automatically screened out for a variety of reasons (self-reported low proficiency/low use of either language, speakers of non-Belgian Dutch). A short proficiency test for both languages followed.

A screen with instructions in CS-mode was presented. The instructions were adapted from Koronkiewicz (2019a) and can be found in appendix a.

Stimuli were pseudo-randomly distributed over two blocks. Presentation of the blocks was randomised, as was presentation of the stimuli within the blocks. Participants were asked to judge a total of 72 sentences. The target stimuli represented roughly one third of the total stimuli. There were 36 sentences used to investigate gender assignment strategies (see chapter 5), and 18 control stimuli, developed according to the guidelines in Koronkiewicz (2019a). Information about and analysis of these control stimuli can be found in appendix c.

After completing all judgments, participants were presented with a few post-test questions about their language use. Participants were asked to indicate their agreement with the following statements on a scale of 1–7. This measured a (self-reported) preference for one of the two languages (14a–b), attitudes towards CS (14c) and mixing behaviour (14c).

(14) a. In everyday conversation, I prefer speaking English.
b. In everyday conversation, I prefer speaking Dutch.
c. People should avoid mixing Dutch and English.
d. In everyday conversation, I mix Dutch and English.

They were also presented with monolingual control items to verify which gender they assigned to each of the nouns included in the set of stimuli, to verify that the variable under investigation was part of the participants' I-language, as recommended by Ebert and Koronkiewicz (2018). Participants that performed unexpectedly – i.e. judging ungrammatical monolingual word orders to be correct – were excluded from analysis.

8.4.3 Predictions

Recall that in Dutch, any adverb can intervene between the verb and the direct object, and that this has been attributed to the high position of Dutch verbs, due to V2. The details of analyses

of V2 are provided in section 7.2. In (9b), repeated here as (15), we can see that the verb is in a high position (C^0).

(15) $[_{CP}$ **Ártemis leest** $[_{IP}$ **leest** $[_{VP}$ **vaak een boek leest** $]$ $]$

This means that whatever drives V2 movement ought to be responsible for the fact that in Dutch all adverbs can intervene between a verb and its direct object. As has been elaborately discussed in chapter 7, however, the traditional account for V2 does not make specific predictions.

While I did not investigate adverbs intervening between the verb and direct object in chapter 7, I did look at V2 word order. I found that it is the finite verb that drives V2 movement, so I would predict that the AC does not hold when the verb is in Dutch, whereas AC violations will result in unacceptability in sentences where the verb is English. This is exactly what Van Dulm (2007) observes for Afrikaans-English CS.

8.4.4 Results

Data was collected in early 2020 using Limesurvey software (LimeSurvey Development Team 2012). A total of 309 participants took the survey, 151 of whom completed it. This high proportion of incomplete questionnaires is due to the automatic screening of participants. Participants who did not fit the profile were screened out. Participants who completed the survey took an average of 29 minutes do to so.

A number of participants were removed manually. Two participants didn't have at least one parent who was Dutch-speaking or English-speaking. Five participants who reported only basic proficiency in English were removed. Participants reporting the use of Dutch (4) and English (5) on a monthly (or less than monthly) basis were excluded too. Participants who scored under 80% proficiency for Dutch (4) and English (17) were also removed. Five participants were excluded because they rated *all* stimuli a 1.

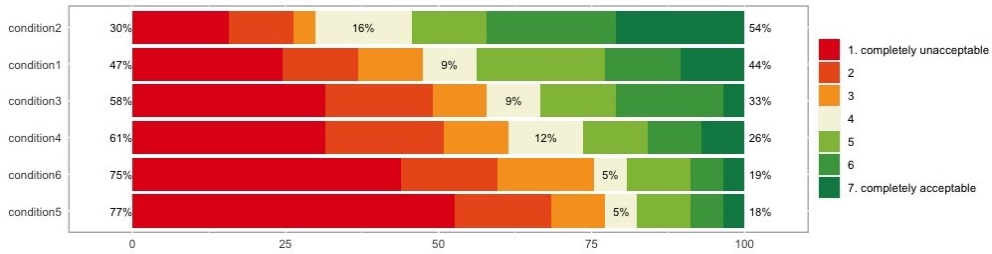


Figure 8.1: Results of the Dutch-English AJT: adverbs

At this point 109 participants remained, 20 of whom were early bilinguals (age of aquisition (AoA) of both English and Dutch below the age of 6). One of those was excluded, since they did not judge AC violations as ungrammatical in monolingual English. This left the responses of 19 participants – three of whom were male, one of whom was non-binary and 15 of whom were female – for analysis. Since each condition was presented 3 times, each condition had 57 (3*19) judgments.

After the manual removal of the unsuitable participants, the data was imported into R-studio (RStudio Team 2020) and prepared for analysis using the **tidyr** and **dplyr** packages (Wickham et al. 2020; Wickham and Henry 2020). Visualisation and summary statistics were provided by the **likert** package (Bryer and Speerschneider 2016). For the statistical modelling, I used the **clmm** function in the **ordinal** package (Christensen 2019). While the output of the **clmm** function does include the Akaike's Information Criterion (AIC)-value, it does not include the AIC_c, which was obtained using the **AICc** function in the **AICcmodavg** package (Mazerolle 2019).

Recall that the AIC_c is appropriate for smaller sample sizes.

Figure 8.1 presents an overview of the ratings per conditions and these are summarised in table 8.2. The conditions are ranked according to acceptability, with the highest ranked condition on top. The table 8.2 shows the mean, standard deviation (StD), and percentage of low (1–3), neutral (4) and high (5–7) ratings.

condition	mean	StD	low	neutral	high
2	4.46	2.15	29.82	15.79	54.39
1	3.68	2.11	47.37	8.77	43.86
3	3.19	2.05	57.89	8.77	33.33
4	3.05	1.99	61.40	12.28	26.32
6	2.53	1.80	75.44	5.26	19.30
5	2.32	1.81	77.19	5.26	17.54

Table 8.2: Summary of the ratings of the adverb stimuli per condition: Dutch-English CS

As we can see, in general the participants didn't really like these kinds of sentences. Even condition 2, with the most positive responses, garnered only just over half (54%) of acceptable responses. This generally low rating is perhaps unsurprising, since all sentences contained an AC violation, which in English would result in ungrammaticality. The high ranking of condition 2 is also not surprising. This condition involved Dutch sentences with only the direct object in English, and according to the existing literature, this is not expected to interfere with grammaticality.

As discussed in section 3.2, I have analysed these results with both parametric and non-parametric tests. For the parametric test, this will be a t-test, while the non-parametric test will be a χ^2 (chi-squared) test.

8.4.4.1 Tests of significance

The first factor I investigated was the impact of the finite verb, and both tests indicated a significant impact on ratings. If we group sentences with a Dutch finite verb together, they are rated significantly higher than those with an English verb (t = 5.43, df = 351.54, p-value < 0.001 |χ^2 = 29.91, df = 6, p-value < 0.001). As for effect size, the t-test indicated a medium-sized effect (Cohen's

d = 0.58, 95% CI [0.37,0.79])[6], while the χ^2-test indicated a large effect (Cramér's V = 0.29, 95% CI [0.15,0.37]).[7] Language of the adverb, on the other hand, had no significant effect on ratings (t = 1.27, df = 357.53, p-value = 0.205 $|\chi^2$ = 5.36, df = 6, p-value = 0.499). Similarly, language of the direct object also did not impact ratings significantly (t = -1.58, df = 356.83, p-value = 0.116 $|\chi^2$ = 4.6862, df = 6, p-value = 0.585).

8.4.4.2 Modelling

I created an ordinal mixed model that predicted the ratings on the basis of the other variables. In the maximal model, I included a random slope and intercept for participant, but only a random intercept for item, as each participant only rated each item once (Judd et al. 2017, p 609). The fixed-effect structure of the maximal model included fixed effects for language of the verb, the direct object and the adverb. I included two extra-linguistic factors: self-reported CS behaviour and attitudes (i.e. agreement with the statements in (14c) and (14d)), as we know from the literature that CS attitudes and behaviour may influence judgments of CS structures.

The AIC_c showed that the maximal model was the model with the best fit. The only non-significant effect was attitude towards CS (i.e. agreement with statement (14c)), but removal resulted in a ΔAIC_c of 0.703, so I left it in the model. Adding an interaction between the effects did not improve the fit of the model. The standard error (SE), p-value, effect size (standardised regression coefficient) with its CI for the factors of the model with the best fit are shown in table 8.3.

The largest effect here is the language of the verb: if it is English, the ratings are significantly lower. The language of the

6 Cohen's d was interpreted using the **interpret_d** function (with the default rule "funder2019") of the **effectsize** package.

7 Cohen (1988, p 222) gives 0.224 for a large effect and 0.134 for a medium effect with df 6.

	SE	p-value	coefficient	95% CI
language of the verb: English	0.458	< 0.001	-2.42	[-3.31, -1.52]
language of the DO: English	0.37	0.038	-0.77	[-1.49, -0.04]
language of the adverb: English	0.38	< 0.001	-1.40	[-2.15, -0.65]
attitudes towards CS	0.224	0.483	-0.30	[-1.13, 0.54]
CS behaviour	0.208	0.013	1.04	[0.22, 1.86]

Table 8.3: Adverbs in Dutch-English CS: Fixed effects for the ordinal mixed model

direct object and adverb have similar effects, though they are smaller in size, comparable to the effect of CS behaviour.

8.5 TESTING ENGLISH-FRENCH

8.5.1 Stimulus-design

Six French and English manner and frequency adverbs were selected.

- QUICKLY
- COMPLETELY
- REGULARLY
- OFTEN
- EASILY
- SOFTLY

- *rapidement*
- *complètement*
- *regulièrement*
- *souvent*
- *facilement*
- *doucement*

The same criteria as described in section 8.4.1 were used to develop the stimuli. An overview of the conditions is presented together with an example in table 8.4. Each condition was presented 4 times, resulting in a total of 24 target stimuli.

	language of the			
	verb	adverb	direct object	example
1	French	English	French	*Le témoin contredisait* COMPLETELY *ta déclaration.* the witness contradicted your statement
2	French	French	English	*Je complétais rapidement* THE SURVEY. I completed quickly
3	French	English	English	*La pianiste remportait* EASILY THE CONTESTS. the pianist won
4	English	French	French	MY LOVE WHISPERED *doucement ces mots.* softly those words
5	English	English	French	THE BRITISH DEFEATED EASILY *leurs adversaires.* their adversaries
6	English	French	English	MR JONES CHANGED *complètement* HIS LOOK. completely

Table 8.4: Adverbs in English-French CS: conditions

8.5.2 Procedure

The procedure was the same as the one described in 8.4.2. However, in this survey, there was a total of 78 sentences, with target stimuli representing roughly one third (24) of the total stimuli. There were 30 sentences used to investigate gender assignment strategies (see chapter 5 for details), and 24 control stimuli, developed according to the guidelines in Koronkiewicz (2019a). Information about and analysis of these control stimuli can be found in appendix C.

8.5.3 Predictions

The analysis presented in Pollock (1989) would straightforwardly predict that the language of T will determine grammaticality. The French verb undergoes V-to-T raising, while the English verb remains low. Consequently, we can predict that in a sentence with an English (lexical) verb, no adverbs should be able to intervene between the verb and its direct object. If the finite verb is French,

on the other hand, a manner or frequency adverb should be able to occur in that position.

Stowell (1981) on the other hand argues that "[i]t may be correct to view the [Romance] manner adverbials as being 'invisible' for the purposes of Case assignment" (p 114). This would mean that regardless of the language of any of the other elements in the sentence, a French adverb (of manner) would be able to intervene between the verb and its direct object, while an English one would not. This is in line with the tentative findings in Koronkiewicz (2019b). While Koronkiewicz said his results were somewhat difficult to interpret, he found that language of the adverb correlated with acceptability. By contrast, the findings in Stadthagen-González et al. (2018) – all the constituents within the domain of the dependency contribute to (un)acceptability – are not expected from the monolingual literature.

8.5.4 Results

Data was collected in early 2020 using Limesurvey software (LimeSurvey Development Team 2012). A total of 154 participants took the survey, 94 of whom completed it. Note that the high proportion of incomplete questionnaires is due to the automatic screening of participants. Participants who did not fit the profile were screened out. Participants who completed the survey took an average of 28,9 minutes to do so.

A number of participants were removed manually. Two participants were removed as they didn't have at least one parent who was English-speaking or French-speaking. Two participants who reported only basic proficiency in English were removed. All participants reported using both English and French on at least a weekly basis. Participants who scored under 80% for French (1) and English (10) proficiency were also removed. One participants was excluded because they rated *all* stimuli a '1'. At this point 78 participants remained, 24 of whom were early bilin-

guals (AoA of both English and French before the age of 6). Two of those were excluded, since they did not judge AC violations as ungrammatical in monolingual English. This left the responses of 22 participants for analysis, 16 of whom were female and 6 of whom were male. As each condition was presented 4 times, each condition had 88 (22*4) judgments.

After the manual removal of the unsuitable participants, the data was imported into R-studio (RStudio Team 2020) analysed with the same packages as were described in 8.4.4.

Figure 8.2 presents an overview of the ratings per conditions and these are summarised in table 8.5. The conditions are ranked according to acceptability, with the highest ranked condition on top. The table 8.2 shows the mean, standard deviation (StD), and percentage of low (1–3), neutral (4) and high (5–7) ratings.

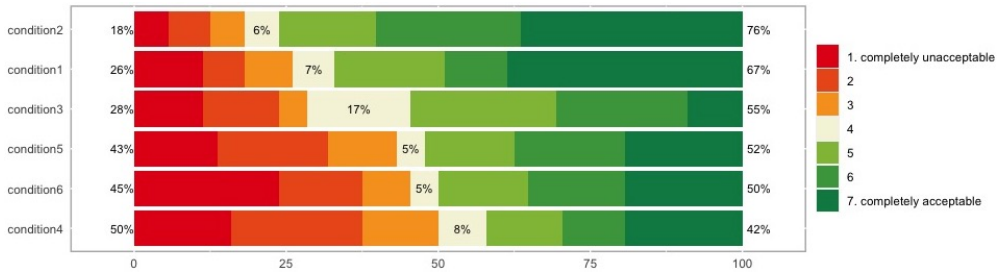


Figure 8.2: Overview of the ratings of the adverb stimuli per condition: English-French cs

Contrary to the Dutch-English data, the participants rather liked these kinds of stimuli. Even the condition with the lowest ratings (condition 4) still had 50% of participants that rated the stimuli a 4 or higher.

8.5.4.1 Tests of significance

So let's delve a bit deeper into the predictions. The finite verb definitely has an impact on ratings. If we group sentences with a French finite verb together, they are rated significantly higher than those with an English verb (t = 5.09, df = 567.59, p-value <

condition	mean	StD	low	neutral	high
2	5.36	1.85	18.18	5.68	76.14
1	4.99	2.13	26.14	6.82	67.05
3	4.31	1.85	28.41	17.05	54.55
5	4.20	2.17	43.18	4.55	52.27
6	3.98	2.31	45.45	4.55	50.00
4	3.88	2.18	50.00	7.95	42.05

Table 8.5: Summary of the ratings of the adverb stimuli per condition: English-French CS

0.001 |χ^2 = 33.46, df = 6, p-value < 0.001). Again, the t-test is more conservative, with a medium effect (Cohen's d = 0.43, 95% CI [0.26, 0.59]), while the χ^2 test indicates a large effect (Cramér's V = 0.24, 95% CI [0.14, 0.31]).

Language of the adverb doesn't significantly impact ratings on either of the two tests (t = 0.5, df = 573.12, p-value = 0.62 |χ^2 = 7.04, df = 6, p-value = 0.318). Language of the direct object, on the other hand, *does* approach significance on the χ^2-test (χ^2 = 12.45, df = 6, p-value = 0.053) but not on the t-test (t = -1.091, df = 572.61, p-value = 0.276). Taking a closer look at the effect size, the t-test gives a Cohen's d of -0,09, with a 95% CI of [-0.26, 0.07], making the effect size between tiny and small. The χ^2 test results in a Cramér's V of 0.15 with a 95% CI between [0.00, 0.20], meaning that the effect is between zero and medium.

8.5.4.2 Modelling

I created an ordinal mixed model that predicted the ratings on the basis of the other variables. In the maximal model, I included a random slope and intercept for participant, but only a random intercept for item. The fixed-effect structure of the maximal model included fixed effects for language of the verb,

	SE	p-value	coefficient	95% CI
language of the verb: French	0.3885	< 0.001	1.66	[0.90, 2.42]
language of the DO: French	0.3136	< 0.001	1.16	[0.55, 1.78]
language of the adverb: French	0.3147	< 0.001	1.41	[0.79, 2.03]
language DO*language adverb	0.4101	0.017	-0.98	[-1.78, -0.18]

Table 8.6: Adverbs in English-French CS: Fixed effects for the ordinal mixed model

the direct object and the adverb. I included two extra-linguistic factors: self-reported CS behaviour and attitudes (i.e. agreement with the statements in (14c) and (14d)).

I then used the AIC_c to trim down the effect structure. The model with the best fit included the effects for language of the verb, direct object and adverb. Adding an interactions between the language of the direct object and adverb also improved the fit of the model. The standard error (SE), p-value, effect size (standardised regression coefficient) with its CI for these factors are shown in table 8.6, together with their significant interactions.

As we can see, language of the verb, direct object and adverb all have a significant effect: if it is French, the sentences are rated higher. The largest effect here is again for the verb, though the difference with the other two is minimal. I looked into the significant interaction using the **emmeans** function in the **emmeans** package (Lenth 2020), with Bonferroni correction for multiple comparisons, and found, unsurprisingly, that when the adverb is English, the stimuli are rated even lower if the direct object is also English (-1.162, SE = 0.314, p < 0.001).

8.6 DISCUSSION AND CONCLUSION

The experiments discussed in the previous sections reveal two main findings. The first one is that, while all constituents in the

predicate had an effect on acceptability, the effect of the verb was the largest for both the Dutch-English and English-French data, though in the English-French data, the difference between the effect of the verb and the other constituents was minimal.

The second main finding is that AC violations are overall better tolerated in English-French CS than in Dutch-English CS. One possible explanation is that the Dutch-English bilinguals are just stricter raters. Thanks to the control stimuli (see appendix c) that were included, we can verify if this is the case. The grammatical control stimuli are rated a 4 or higher by 56-73% of the Dutch-English bilinguals, while 85-91% of the English-French bilinguals rate these sentences a 4 or more. This indicates that it is indeed the case that the Dutch-English bilinguals (in this study) were harsher raters.

However, it does not seem like this is the full story. In the figure in the appendix that presents the results from the English-French bilinguals (figure c.3), there is a clear break between conditions 2 (the lowest rated grammatical condition) and 6 (the highest rated ungrammatical condition). The conditions presented in this chapter are all rated better than the best rated ungrammatical control condition. So it seems that the full story is a combination of Dutch-English bilinguals being stricter raters *and* the English-French bilinguals being unexpectedly permissive for AC violations in sentences containing English elements.

To my knowledge, this is only the fourth study which focuses on adverb placement in CS. All three previous studies deal with language pairs that are typologically similar to the pairs under investigation here. All involve a language with a strict AC (English) and one language in which AC violations don't (always) result in ungrammaticality (Spanish: Koronkiewicz 2019a; Stadthagen-González et al. 2018 and Afrikaans: Van Dulm 2007). All three of these previous studies have different, maybe even contradictory, findings. To me, this is quite surprising, as there seems to be a consensus in the (monolingual) literature on how exactly the

differences in adverb placement arise in languages such as Dutch, French and English.

The findings of the present study are most consistent with the results in Stadthagen-González et al. (2018): all constituents in the clause have some effect on acceptability. This finding highlights once more the important contribution of CS data to debates within linguistic theory.

Methodologically, this chapter has shown once more that the "pragmatic sanction" invoked by Stevens (1951, p. 26) (as cited in Bross 2019a, p. 47) is valid. In all cases the parametric tests I performed gave the same result as the non-parametric counterparts. Additionally, the statistical modelling provided an insight that the other tests could not: the models showed that there were multiple linguistic factors that had a statistically significant effect on grammaticality, which was not detected by the classic tests of significance.

CONCLUSION

Over the last two decades, the realisation has grown that linguistic data from bi- and multilingual individuals can provide a spotlight to uncover observations and generalisations about linguistic structure that would otherwise remain in the shadows. In this dissertation, I have examined some of the the implications of using code-switching data to inform linguistic theory.

I have done so by means of several case studies of English-French and Dutch-English code-switching in two domains: grammatical gender and word order. In this chapter, I will summarise my conclusions for each of the studies that were undertaken within the context of this dissertation. I will also discuss the methodological contributions of the dissertation as well as its limitations, and outline some suggestions for future research.

GENDER AGREEMENT

The first set of research questions concerned the domain of grammatical gender gender. In the first chapter, I delved into the theory of grammatical gender. While I broadly adopted the approach developed by Kramer (2015), I made several changes. Firstly, I proposed that for some roots, conceptual gender of their potential referent is represented in the lexical meaning of those roots. Since I make this assumption, I no longer need Kramer's distinction between interpretable and uninterpretable gender features (on *n)* and the inventory of *n*s required to account for the grammatical gender in Dutch and French is reduced.

In the chapter in which I discuss the contribution of code-switching data to the gender literature, I investigated gender-

agreement strategies in Dutch-English and English-French CS. I found that for both language pairs, sentences with a determiner marked for grammatical gender had a lower mean rating than those that had a determiner not marked for grammatical gender for all types of bilinguals. This could be an indication of gender-avoidance.

As for gender agreement, the results were quite interesting. Recall that there is a robustly observed pattern in the literature that early bilinguals tend to go for the default gender strategy, whereas late bilinguals tend to assign code-switched nouns the gender of the translational equivalent. This pattern was observed for neither language pairs under investigation in this chapter.

Overall, the Dutch-English bilinguals, both early and late, had a preference for default gender, regardless of their AoA of English. The default gender seemed to be the feminine, and not neuter as predicted in the literature, nor masculine, which was the second likely candidate for default. I argued that this surprising result could be due to influence of Standard Dutch, where the common gender (whose determiners are formally identical to those of the feminine in Belgian Dutch) is the default.

In contrast to the Dutch-English bilinguals, both the early and late English-French bilinguals all had a preference for the analogical gender strategy.

WORD ORDER

Within the domain of word order I looked at three different domains: verb-second word order, word order in the subordinate clause and adverb position. For verb-second word order I argued that there seems to be an issue with the traditional Generative account for verb second word order when it comes to incorporating CS data. This account seems to make no predictions for bilingual data.

In this dissertation, I provided evidence that in Dutch-English code-switched sentences, word order is aligned with the finite verb in both main and subordinate clauses. For the subordinate clause, this is further evidence for accounts attributing verb-object order to T^0. This speaks in favour of accounts where the word order in subordinate clauses is determined by T^0, rather than C^0. For V2 constructions in the main clause, this seems to cast some doubt on the mainstream Generative account, which ascribes V2 word order to C^0. The V2 data in this chapter indicate that there is a close connection between T^0 and C^0. I also argued that none of the mechanisms proposed in the code-switching literature provide an adequate solution to account for this close connection.

A possible mechanism that *could* account for the data presented in this dissertation is the mechanism of Feature Inheritance of Tense features from C^0 to T^0. However, I remain unconvinced that this mechanism straightforwardly applies in situations where C^0 is unlexicalised. A proposal along the lines of the one put forward by Gallego (2014), in which C^0 and T^0 are the same lexical unit, seems initially more promising, but it remains an open question whether CS between C^0 and T^0 is compatible with that proposal. Hence, unless FI can be robustly shown to apply in main clauses (with an unlexicalised C^0), the data presented in this dissertation present a challenge to Generative approaches to verb second.

The second set of research questions in the domain of word order concerns adverb position. While the monolingual literature seems pretty much in agreement on what drives the difference in the position of adverbs between Dutch, English and French, the recent CS literature had cast some doubt on the common assumptions.

The experiments discussed in the chapter that concerned adverb position reveal two main findings. The first one is that, while all constituents in the predicate had an effect on acceptability

for both language pairs. For Dutch-English, the effect of the verb was much larger than the other effects, whereas for the English-French data the effect of the verb was comparable to the effects of the direct object and adverb.

The second main finding is that AC violations are overall better tolerated in English-French CS than in Dutch-English CS. In fact, the English-French bilinguals were unexpectedly permissive for the placement of an adverb between a verb and its direct object in sentences containing English elements.

METHODOLOGICAL CONTRIBUTION

In section 3.2, I introduced the methodological issue of which statistical methods are most appropriate for the analysis of Likert-type data. While some researchers insist on the use of non-parametric tests for the analysis of Likert items, others argue that parametric tests are so robust that they can be used to analyse them. In this dissertation, I have analysed my data with both types of tests and found overall very little difference in terms of the statistical significance of the contrasts I examined. Overall, it seems like it is indeed the case that using parametric tests on ordinal data does not result in an increased rate of type I errors.

In addition, I also used ordinal mixed models for my data analysis. This statistical modelling gave insight into the data that the traditional tests did not, particularly in the case of the gender agreement data. For almost all data-sets, the models showed that extra-linguistic factors *can* influence acceptability judgments, but if they do, that linguistic factors have a larger effect.

LIMITATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

The biggest limitation of the study is that it is based entirely on data collected through an acceptability judgment task. While section 3.1 shows that acceptability judgment tasks are a valuable and valid method to collect linguistic data, and that they can also be applied in a code-switching context, they do not tell the full story. In order to gain a full understanding of humans' linguistic capacities, every kind of datatype has its place. As Fillmore (1992, p 35) rightly said:

> I have two main observations to make. The first is that I don't think there can be any corpora, however large, that contain information about all of the areas of English lexicon and grammar that I want to explore; all that I have seen are inadequate. The second observation is that every corpus that I've had a chance to examine, however small, has taught me facts that I couldn't imagine finding out about in any other way.

Comparing the results from the studies collected in this dissertation with data from corpora could provide additional insights. In particular the lack of evidence for a mixed-DP asymmetry in both Dutch-English and English-French CS should be verified with corpus data, as it is consistently observed in bilingual corpora involving other pairs of languages, and never in AJTs. Psycholinguistic and neurocognitive studies may also provide further understanding of the processes involved in CS.

Another limitation concerns the study on grammatical gender agreement in Dutch-English CS. It is very likely that a prescriptive bias against Tussentaal influenced the results of the study presented in section 5.2. While some measures to avoid such a bias were taken, it is apparent these weren't sufficient. A follow-up study, with additional focus on accommodating Tussentaal, could perhaps help eliminate this confound.

As regards the placement of the adverb, the experiments in chapter 8 presented a puzzling finding: why are AC violations tolerated so much better by English-French bilinguals compared to the Dutch-English bilinguals?

I think analysing CS data has brought some questions to the fore that we are not in a position to ask when adhering to a lexicalist framework. Whereas one of the questions in the field has been whether CS within a word is allowed or not, it seems clear that it *is* allowed, but that there are rules governing such word-internal switches. A DM-like framework allows us to address these questions in a way that lexicalist approaches do not. However, as I pointed out in section 2.5, there are many details to be worked out there as well. The relevance of a Distributed Morphology (DM) approach was highlighted in the chapters that concerned word formation (chapters 4 and 5) in particular.

FINAL WORDS

Well, we made it to the end. I hope you enjoyed reading this dissertation as much as I enjoyed (what went into) writing it. If there's one thing that I have taken away from more than seven years as a CS researcher is that CS is not fundamentally different from other language use. The same motivations, principles and restrictions that underlie "monolingual" language use can and should be applied to code-switching. And just like for monolingual language use, to gain a full understanding of a multi-faceted phenomenon like CS, a combination of different approaches and methods is indispensable.

Part IV

APPENDIX

a | INSTRUCTIONS

DATA COLLECTION 2017

In dit deel wordt je gevraagd zinnen te beoordelen. Please indicate to what degree you find them acceptable. Met "acceptabel" bedoelen we in welke mate je de structuur van de zin goed vindt. You'll notice dat alle zinnen zowel Engels als Nederlands bevatten. We zijn geïnteresseerd in jouw mening about the interaction between these two languages.

DATA COLLECTION 2020

The instructions that were presented to the participants in this round of data-collection were adapted from Koronkiewicz (2019a).

Dutch-English

In this study we are interested in finding out hoe tweetaligen afwisselen tussen Nederlands en Engels in een tweetalige conversatie. Dit noemen we "code-switching". Code-switching is a form of linguistic expression like any other and, therefore, it is subject to rules and restrictions like any other. De regels en restricties die we hier bedoelen hebben niets te maken met "correct taalgebruik" zoals aangeleerd wordt op school, but rather with the linguistic structures that speakers have in their minds. Let us give you an example in English:

(1) There is likely that John likes Mary.

We will ask you to rate sentences from 'completely unacceptable' to 'completely acceptable'. The sentence in (1) would be 'completely unacceptable' because native speakers of English find this sentence very strange. It is perfectly understandable but there is something about its structure that sounds un-English. The following sentence – daarentegen – sounds completely fine:

(2) There is someone in the garden.

Een voorbeeld in het Nederlands:

(3) Deze film Marie vond niet leuk.

Deze zin is niet acceptabel. Hoewel hij perfect verstaanbaar is, is het geen Nederlandse zin. De zin in (4) daarentegen is perfect normaal.

(4) Jan kocht een zak appels.

Hier hebben we two examples in code-switching:

(5) Mijn moeder gave haar a house.

De meeste Nederlands/Engels tweetaligen who practice code-switching regularly or occasionally agree that sentence (5) sounds strange.

Sentence (6), on the other hand, sounds completely fine voor de meeste tweetaligen.

(6) Ik vind that John is a bit weird.

In the survey, you might find that some sentences are neither perfect nor totally awful. Daarom kan je gebruik maken van een schaal van 1 tot 7. Neem dan nu de volgende zin.

(7) I know my neighbour chewed gum yesterday.

This sentence may sound odd because it doesn't seem like something anyone would ever say or care about. The question, however, is of dit een mogelijke Engelse zin is. Dat is het geval, dus de zin is volledig acceptabel.

Now, take a moment to read the following sentence:

(8) He surprise that no one called yesterday.

This sentence is comprehensible and the situation seems plausible (bijvoorbeeld: gisteren verwachtte hij een telefoontje, maar uitendelijk heeft niemand hem gebeld), but the sentence is just not English, and therefore it is completely unacceptable. Bij het beoordelen van de zinnen, the question, then, is whether the sentence is possible in English, Dutch or code- switching, even if you don't know why anyone would actually say the sentence.

Een laatste opmerking: We are also interested in linguistic markers of "Tussentaal". Tussentaal is een informeel register dat gebruikt wordt door sprekers van het Nederlands in Vlaanderen. Hoewel sprekers van het Belgisch Nederlands vaak wordt aangeraden Tussentaal te vermijden, zien wij Tussentaal als een volwaardige taalvariant, much like code- switching. Als je zo een tussentaalmarkeerder tegenkomt in de enquête, moet je de zin enkel beoordelen als "volledig onacceptabel" als de zin niet zou kunnen voorkomen in Tussentaal.

(9) Ge zijt zot geworde.

Zin (9) is zeker geen Standaard Nederlands, but sentences like it are certainly used by speakers of Tussentaal and it is therefore considered completely acceptable. Sentence (10) on the other hand, sounds very odd to Tussentaal speakers, and is considered completely unacceptable.

(10) Marie is mijne vriendin.

English-French

In this study we are interested in finding out comment les personnes bilingues alternent entre le français et l'anglais dans une conversation bilingue. En anglais, cela s'appelle le « code-switching ».

Code-switching is a form of linguistic expression like any other and, therefore, it is subject to rules and restrictions like any other. Ces règles et restrictions n'ont rien à voir avec « le bon usage » de la langue tel qu'il est enseigné à l'école, but rather with the linguistic structures that speakers have in their minds. Let us give you an example in English:

(1) There is likely that John likes Mary.

We will ask you to rate sentences from 'completely unacceptable' to 'completely acceptable'. The sentence in (1) would be 'completely unacceptable' because native speakers of English find this sentence very strange. It is perfectly understandable but there is something about its structure that sounds un-English. The following sentence - en revanche - sounds completely fine :

(2) There is someone in the garden.

Un exemple en français :

(3) C'est trois ans que je parle le français.

Cette phrase n'est pas acceptable. Bien qu'elle soit compréhensible, ce n'est pas une phrase du français. La phrase (4) en revanche est parfaitement normale.

(4) Jean a acheté un sac de patates.

Voici maintenant deux exemples avec code-switching:

(5) Ma mère gave lui a house.

La plupart des bilingues français-anglais who practice code-switching regularly or occasionally agree that sentence (5) sounds strange.

Sentence (6), on the other hand, sounds completely fine pour la majorité des bilingues.

(6) Mes nièces se sont baignées in the pool for three hours.

In the survey, you might find that some sentences are neither perfect nor totally awful. C'est pourquoi nous vous proposons une échelle de 1 à 7. Prenons la phrase suivante.

(7) I know my neighbor chewed gum yesterday.

This sentence may sound odd because it doesn't seem like something anyone would ever say or care about. The question, however, is whether this is une phrase possible en anglais. Comme c'est le cas, cette phrase est tout à fait acceptable. Now, take a moment to rate the following sentence:

(8) He surprise that no one called yesterday.

This sentence is comprehensible and the situation seems plausible (par exemple : elle attendait un coup de fil hier, mais finalement personne n'a appelé), but the sentence is just not English, and therefore it is completely unacceptable. Chaque fois que vous devrez juger une phrase, the question, then, is whether the sentence is possible in English, French or code- switching, even if you don't know why anyone would actually say the sentence.

b | STIMULI

item	condition	lg D	lg N	gen D	(A)Gen N	num	
NLGen1	1	Dutch	English	M	M	sg	De tandarts trok den tooth.
NLGen2	2	Dutch	English	M	F	sg	Mieke droeg den dress.
NLGen3	3	Dutch	English	M	N	sg	De rotsklimmer gebruikte den chalk.
NLGen4	4	Dutch	English	F	M	sg	Henk verbeterde de bike.
NLGen5	5	Dutch	English	F	F	sg	Mijn zus schilderde de drawer.
NLGen6	6	Dutch	English	F	N	sg	Zij groeven de hole.
NLGen7	7	Dutch	English	N	M	sg	De kinderen haakten het scarf.
NLGen8	8	Dutch	English	N	F	sg	Haar ouders betaalden het rent.
NLGen9	9	Dutch	English	N	N	sg	Marie repareerde het roof.
NLGen10	1	Dutch	English	M	M	sg	Mijn collega kocht den chair.
NLGen11	2	Dutch	English	M	F	sg	Mijn ouders vulden den box.
NLGen12	3	Dutch	English	M	N	sg	De politiedetectives vonden den body.
NLGen13	4	Dutch	English	F	M	sg	Mijn partner gaf me de key.
NLGen14	5	Dutch	English	F	F	sg	Mijn vader kreeg de flower.
NLGen15	6	Dutch	English	F	N	sg	Deze kamer heeft de window.
NLGen16	7	Dutch	English	N	M	sg	De kleermaker ontwierp het skirt.
NLGen17	8	Dutch	English	N	F	sg	De ballerina huurde het room.
NLGen18	9	Dutch	English	N	N	sg	De koningin kreeg het watch.
NLGen19	10	Dutch	English	PLUR	PLUR	plural	De babysitter kietelde de bellies.
NLGen20	10	Dutch	English	PLUR	PLUR	plural	De advocaten wonnen de trials.
NLGen21	10	Dutch	English	PLUR	PLUR	plural	De bokser ontweek de blows.
NLGen22	10	Dutch	English	PLUR	PLUR	plural	De kinderen vulden de days.
NLGen23	10	Dutch	English	PLUR	PLUR	plural	De biologen verzamelden de barks.
NLGen24	10	Dutch	English	PLUR	PLUR	plural	De schrijnwerker maakte de desks.
NLGen25	11	English	Dutch	–	M	sg	My neighbour knitted the sjaal.
NLGen26	11	English	Dutch	–	M	sg	A petty thief stole the fiets.
NLGen27	11	English	Dutch	–	F	sg	My father hid the muts.
NLGen28	11	English	Dutch	–	F	sg	We filled the schuif.
NLGen29	11	English	Dutch	–	N	sg	The undertaker dug the gat.
NLGen30	11	English	Dutch	–	N	sg	Mary repaired the dak.
NLGen31	12	English	Dutch	–	–	plural	They used the stoelen.
NLGen32	12	English	Dutch	–	–	plural	We like the rokken.
NLGen33	12	English	Dutch	–	–	plural	My mother bought the bloemen.
NLGen34	12	English	Dutch	–	–	plural	Mij friends needed the jurken.
NLGen35	12	English	Dutch	–	–	plural	We visited the landen.
NLGen36	12	English	Dutch	–	–	plural	They cleaned the ramen.

Table b.1: List of Dutch-English stimuli for the gender agreement survey

item	condition	lg D	lg N	gen D	(A)Gen N	num	
FRGen1	1	French	English	M	M	sg	Le bijoutier réparait le necklace.
FRGen2	1	French	English	M	M	sg	Marie arrangeait le roof.
FRGen3	1	French	English	M	M	sg	Le client choisissait le bag.
FRGen4	2	French	English	M	F	sg	Le tailleur dessinait le dress.
FRGen5	2	French	English	M	F	sg	Jean se massait le leg.
FRGen6	2	French	English	M	F	sg	La danseuse louait le room.
FRGen7	3	French	English	F	M	sg	Le jardinier arrosait la tree.
FRGen8	3	French	English	F	M	sg	Les élèves lisaient la book.
FRGen9	3	French	English	F	M	sg	Les invitées buvaient la coffee.
FRGen10	4	French	English	F	F	sg	Ma mère regardait la moon.
FRGen11	4	French	English	F	F	sg	L'homme remplissait la box.
FRGen12	4	French	English	F	F	sg	Les voyageurs examinaient la map.
FRGen13	5	French	English	PLUR	M	plural	Jean-Philippe se lavait les feet.
FRGen14	5	French	English	PLUR	M	plural	Le fossoyeur creusait les holes.
FRGen15	5	French	English	PLUR	M	plural	Le concièrge ballayait les sols.
FRGen16	6	French	English	PLUR	F	plural	Les filles comptaient les mistakes.
FRGen17	6	French	English	PLUR	F	plural	Elles aménageaient les wardrobes.
FRGen18	6	French	English	PLUR	F	plural	Mon frère cachait les keys.
FRGen19	7	English	French	—	M	sg	Those thieves stole the vélo.
FRGen20	7	English	French	—	M	sg	Mr Smith fixed the toît.
FRGen21	7	English	French	—	M	sg	We filled the the tiroir.
FRGen22	8	English	French	—	F	sg	His sister opened the boîte.
FRGen23	8	English	French	—	F	sg	The astronomer observed the lune.
FRGen24	8	English	French	—	F	sg	My mother bought the jupe.
FRGen25	9	English	French	—	M	plural	The farmer planted the arbres.
FRGen26	9	English	French	—	M	plural	The worker painted the murs.
FRGen27	9	English	French	—	M	plural	The students browsed the livres.
FRGen28	10	English	French	—	F	plural	The physiotherapist massaged the jambes.
FRGen29	10	English	French	—	F	plural	The florist arranged the fleurs.
FRGen30	10	English	French	—	F	plural	They cleaned the fenêtres.

Table b.2: List of English-French stimuli for the gender agreement survey

item	condition	word order	Lg of Fronted Constituent	finite V	lg of object	
A11	1	V2	Dutch	English	Dutch	In de tuin caught the cat een muis.
A12	2	no V2	Dutch	English	Dutch	Gisteren the boy ate een rijsttaart.
A13	3	V2	English	Dutch	Dutch	Friday belde ik mijn moeder.
A14	6	no V2	English	English	Dutch	In March Jean celebrates haar verjaardag.
A15	7	V2	Dutch	Dutch	English	Wekelijks koken mijn ouders an Asian dish.
A16	8	no V2	Dutch	Dutch	English	's Ochtends mijn buren maken noise.
A17	9	V2	Dutch	English	English	Op de trein saw they a dog.
A18	10	no V2	Dutch	English	English	In de slaapkamer I found the keys.
A19	12	no V2	English	Dutch	English	Every morning de priester zegent the holy water
B11	1	V2	Dutch	English	Dutch	In de klas teaches the teacher wiskunde.
B12	2	no V2	Dutch	English	Dutch	Vandaag they watched een actiefilm.
B13	4	no V2	English	Dutch	Dutch	In the evening hij kuste zijn dochter.
B14	5	V2	English	English	Dutch	Every week cook my parents een Oosters gerecht.
B15	6	no V2	English	English	Dutch	In the book Thelma draws een schets.
B16	8	no V2	Dutch	Dutch	English	In maart Janne viert her birthday.
B17	9	V2	Dutch	English	English	Vrijdag called I my mother.
B18	11	V2	English	Dutch	English	In the garden vangt de kat a mouse.
B19	12	no V2	English	Dutch	English	Yesterday de jongen maakte rice pudding.
C11	1	V2	Dutch	English	Dutch	In die winkel buys the dancer haar jurken.
C12	3	V2	English	Dutch	Dutch	Next year bezoeken wij het Rijksmuseum.
C13	4	no V2	English	Dutch	Dutch	On Sundays onze zoon speelt voetbal.
C14	5	V2	English	English	Dutch	Sometimes forget students hun huiswerk.
C15	7	V2	Dutch	Dutch	English	In Oostenrijk ontmoette mijn vader my mother.
C16	8	no V2	Dutch	Dutch	English	Op het blad Tine tekende a sketch.
C17	10	no V2	Dutch	English	English	's Avonds he kissed his daughter.
C18	11	V2	English	Dutch	English	In the classroom geeft de juf mathematics.
C19	12	no V2	English	Dutch	English	Today ze bekeken an action movie.
D11	2	no V2	Dutch	English	Dutch	Elke ochtend the priest blesses het wijwater.
D12	3	V2	English	Dutch	Dutch	In the train zagen ze een hond.
D13	4	no V2	English	Dutch	Dutch	In the bedroom ik vond de sleutels.
D14	6	no V2	English	English	Dutch	In the morning my neighbours make lawaai.
D15	7	V2	Dutch	Dutch	English	Soms vergeten studenten their homework.
D16	9	V2	Dutch	English	English	Volgend jaar visit we the British museum.
D17	10	no V2	Dutch	English	English	Op zondag our son plays football.
D18	11	V2	English	Dutch	English	In the shop koopt de danseres her dresses.
D19	5	V2	English	English	Dutch	In Austria met my father mijn moeder.

Table b.3: List of main clause stimuli for the verb second survey

item	condition	lg finite V embedded clause	comp switch	embedded clause word order	
A1	A	Dutch	before	SVO	My brother said dat Piet kust Sarah.
A2	B	Dutch	before	SOV	Archie guessed dat de ober Jan vermoordde.
A3	C	Dutch	after	SVO	She said that de kinderen gooiden stenen.
A4	D	Dutch	after	SOV	Harry wanted that de hond de weg vond.
A5	E	English	before	SVO	Marie dacht that cats eat fish.
A6	F	English	before	SOV	Je moeder geloofde that Heather Myrtle knew.
A7	G	English	after	SVO	Cedric dacht dat she saw him.
A8	H	English	after	SOV	Linde geloofde dat my neighbour the truth told.
A9	A	Dutch	before	SVO	My cousin dreamt dat een hond beet haar zoon.
A10	B	Dutch	before	SOV	Cleo expected dat Jelle een fiets kocht.
B1	A	Dutch	before	SVO	Mary thought dat de kat eet vogels.
B2	B	Dutch	before	SOV	Your mother believed dat je Mieke kende.
B3	C	Dutch	after	SVO	Charles thought that ze zag hem.
B4	D	Dutch	after	SOV	Gertie believed that hij de waarheid sprak.
B5	E	English	before	SVO	Hij hoopt that they won the match.
B6	F	English	before	SOV	Ze vreesde that she the keys forgot.
B7	H	English	after	SVO	Felix hoopte dat the women knew the solution.
B8	H	English	after	SOV	Hij vermoedde dat his mother a pie baked.
B9	C	Dutch	after	SVO	Ivy dreamt that haar broer bouwde een huis.
B10	D	Dutch	after	SOV	The farmer expected that de koe melk gaf
C1	A	Dutch	before	SVO	He hopes dat ze wonnen de wedstrijd.
C2	B	Dutch	before	SOV	She feared dat ze de sleutels vergat.
C3	C	Dutch	after	SVO	Felix hoped that zijn zus vond een oplossing.
C4	D	Dutch	after	SOV	He suspected that zijn moeder een taart bakte.
C5	E	English	before	SVO	Mijn broer zei that John kisses Sarah
C6	E	English	before	SVO	Sven wist that Rosalind surprised her husband.
C7	F	English	before	SOV	Mijn vader zag that the ball the wall touched.
C8	G	English	after	SVO	De student wist dat the teacher gave the answer.
C9	H	English	after	SOV	De advocaat wist dat the suspect the gold stole.
C10	H	English	after	SOV	Hans wou dat his friend the way found.
D1	A	Dutch	before	SVO	George knew dat Janne verraste haar man.
D2	B	Dutch	before	SOV	My father saw dat de bal de muur raakte.
D3	C	Dutch	after	SVO	The student knew that de leraar gaf het antwoord.
D4	D	Dutch	after	SOV	The lawyer knew that de verdachte het goud stal.
D5	E	English	before	SVO	Mijn nicht droomde that a dog bit her son.
D6	F	English	before	SOV	Ludo raadde that the butler James killed.
D7	F	English	before	SOV	Frieda verwachtte that Sue a bike bought.
D8	G	English	after	SVO	Ze zei dat the children threw the rocks.
D9	G	English	after	SVO	Ine droomde dat her brother built a house.
D10	H	English	after	SOV	De boer verwachtte dat the cow milk gave.

Table b.4: List of subordinate clause stimuli for the verb second survey

item	condition	lg S+V	lg adv	lg obj	
NLAdv1	1	Dutch	English	Dutch	Jozefien at quickly twintig druiven.
NLAdv2	1	Dutch	English	Dutch	Inge poetste daily haar tanden.
NLAdv3	1	Dutch	English	Dutch	Jan bezoekt regularly de dierentuin.
NLAdv4	2	Dutch	Dutch	English	Ik verstopte snel the easter eggs.
NLAdv5	2	Dutch	Dutch	English	Johannes bekeek dagelijks a movie.
NLAdv6	2	Dutch	Dutch	English	Wij zoeken regelmatig a job.
NLAdv7	3	Dutch	English	English	Mijn vader poetste thoroughly his glasses.
NLAdv8	3	Dutch	English	English	Die mannen bouwden patiently a house of cards.
NLAdv9	3	Dutch	English	English	De buurman bakte often cookies.
NLAdv10	4	English	Dutch	Dutch	My mother examined grondig de krant.
NLAdv11	4	English	Dutch	Dutch	We completed geduldig de puzzel.
NLAdv12	4	English	Dutch	Dutch	Edith buys vaak tweedehandskledij.
NLAdv13	5	English	English	Dutch	The teacher spotted easily de fout.
NLAdv14	5	English	English	Dutch	My brother dropped loudly de borden.
NLAdv15	5	English	English	Dutch	The children sang softly de liedjes.
NLAdv16	6	English	Dutch	English	My husband found gemakkelijk a present.
NLAdv17	6	English	Dutch	English	Richard yelled luid a curse.
NLAdv18	6	English	Dutch	English	My sister whispered zachtjes those words.

Table b.5: List of Dutch-English stimuli for the adverb survey

item	condition	lg S+V	lg adv	lg obj	
FRAdv1	1	French	English	French	Cet élève comprenait quickly ses erreurs.
FRAdv2	1	French	English	French	Le témoin contredisait completely ta déclaration.
FRAdv3	1	French	English	French	Sa maman appelait regularly l'hôpital.
FRAdv4	1	French	English	French	Mon frère achetait often des BDs.
FRAdv5	3	French	English	English	La pianiste remportait easily the contests.
FRAdv6	3	French	English	English	Sa soeur fermait softly the door.
FRAdv7	3	French	English	English	Le proviseur rejetait quickly their proposals.
FRAdv8	3	French	English	English	L'architecte rénovait completely the building.
FRAdv9	2	French	French	English	Je complétais rapidement the survey.
FRAdv10	2	French	French	English	Mon père oubliait complètement his worries.
FRAdv11	2	French	French	English	Son ami cherchait régulièrement a carpenter.
FRAdv12	2	French	French	English	Le boulanger vendait souvent those pastries.
FRAdv13	4	English	English	French	The writer expressed regularly ses opinions.
FRAdv14	4	English	English	French	Harry baked often des biscuits.
FRAdv15	4	English	English	French	The British defeated easily leurs adversaires.
FRAdv16	4	English	English	French	The actor sang softly la chanson.
FRAdv17	5	English	French	French	My neighbour visited régulièrement des musées.
FRAdv18	5	English	French	French	Ms. Smith bought souvent des vêtements.
FRAdv19	5	English	French	French	The kids found easily les cadeaux.
FRAdv20	5	English	French	French	My love whispered doucement ces mots.
FRAdv21	6	English	French	English	We finished rapidement our dinner.
FRAdv22	6	English	French	English	Mr Jones changed complètement his look.
FRAdv23	6	English	French	English	My friend completed facilement the marathon.
FRAdv24	6	English	French	English	The boy kissed doucement the teddy bear.

Table b.6: List of English-French stimuli for the adverb survey

C | CONTROL STIMULI

> Enfin, tous les énoncés ayant uniquement le pronom à l'initiale dans une autre langue que le reste de la phrase m'ont paru très étranges.
>
> — participant 10

The studies described in chapters 5 and 8 included CS control stimuli, as recommended by Koronkiewicz (2019a). He advocates for the inclusion of code-switched control stimuli with known acceptability, as they can 'establish a baseline comparison of acceptability' (p 1). He proposes several types of switches: complex-sentence switches, subject–predicate switches, direct-object switches, pronoun switches, and present–perfect switches. Based on the literature, the first three are expected to be acceptable, while the latter two are expected to be ungrammatical. The inclusion of such stimuli allows for the exclusion or isolation of participants who fail to make the expected distinction between the control stimuli.

In this appendix, I elaborate on the types of control stimuli that were included in the surveys. I will also provide an overview of the results for these stimuli.

C.1 DUTCH-ENGLISH CONTROLS

In table C.1, you can see what kind of stimuli were included. Conditions 1–3 are expected to be grammatical, while conditions 4–6 are expected to be ungrammatical. I used sentences with word-order transfer instead of Koronkiewicz's "present-perfect" switches, as I was unsure if these would be ruled our for Dutch-English CS.

	type of switch	example
1	full DP subject	**De buurvrouw** CLEANS HER PORCH EVERY DAY.
2	full DP DO	**Jan en ik drinken** A FEW BEERS.
3	subordinate clause	**Hij was te laat** BECAUSE HE MISSED THE TRAIN.
4	transfer: word order	**Marie wou** THAT I HER HELPED.
5	pronominal subject (Eng)	THEY **geloven echt dat ze de beste zijn.**
6	pronominal subject (Dutch)	**Hij** SEEMED EXTREMELY BORED TODAY.

Table C.1: Dutch-English CS: conditions for the control stimuli

The results of the survey are shown in figure C.1. We can see that the early bilinguals perform as expected: they rate conditions 1, 2 and 3 better than conditions 4, 5 and 6. The late bilinguals (figure C.2) don't seem to rate conditions 2 and 6 differently, and both are judged to be at least somewhat acceptable by more than half of the participants. For condition 2 this is expected, but for condition 6 it is most definitely not.

Also intriguing is the difference in rating between conditions 5 and 6…

This result strengthens the argument for including only early bilinguals in the analysis of the results of chapter 8.

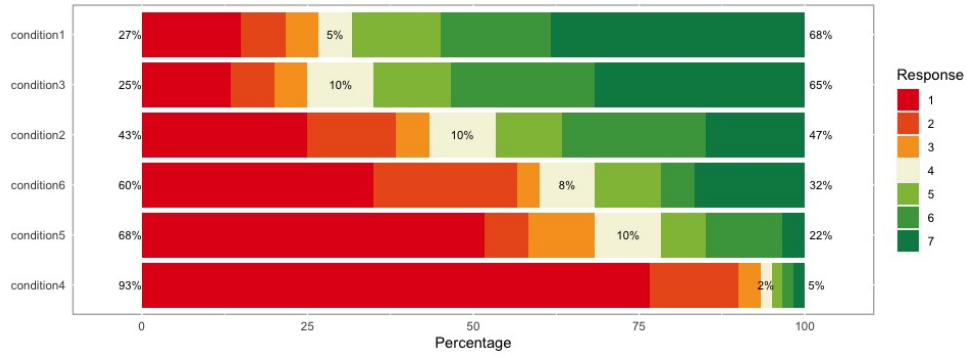


Figure c.1: Dutch-English controls: early bilinguals

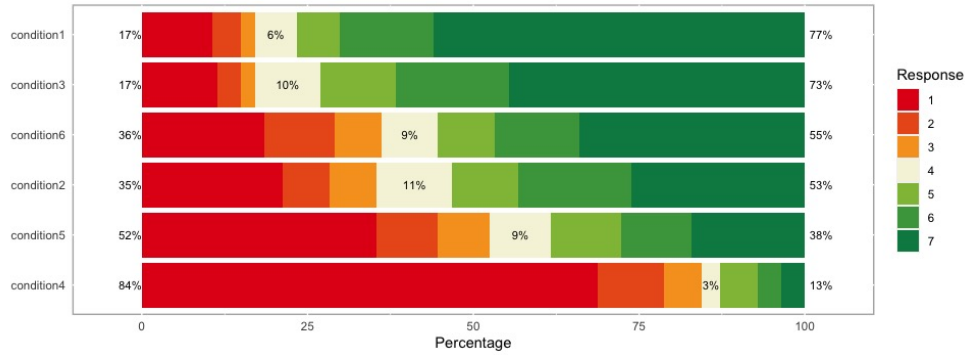


Figure c.2: Dutch-English controls: late bilinguals

C.2 ENGLISH-FRENCH CONTROLS

In table c.1, you can see what kind of stimuli were included. Conditions 1–3 are expected to be grammatical, while conditions 4–6 are expected to be ungrammatical. Again, I used sentences with word-order transfer instead of Koronkiewicz's "present-perfect" switches, as I was unsure if these would be ruled our for English-French CS.

As is shown in figures c.3 and c.4, both the early and late bilinguals perform as expected.

	type of switch	example
1	full DP subject	*Ma grande sœur* REALLY LIKES TO TEASE ME.
2	full DP DO	*Jean et ses amis boivent* A FEW BEERS.
3	subordinate clause	*Il est arrivé si tard* THAT HE MISSED THE TRAIN.
4	L1 transfer: word order	*Son frêre regardait* A MOVIE SCARY
5	pronominal subject (Eng)	WE *allons au marché tous les dimanches.*
6	pronominal subject (Fr)	*Vous* KNOW HOW MUCH I LIKE CHOCOLATE.

Table c.2: English-French CS: conditions for the control stimuli

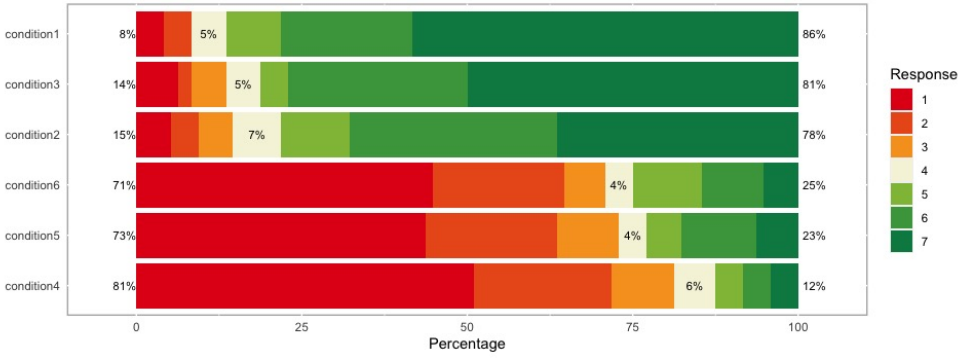


Figure c.3: English-French controls: early bilinguals

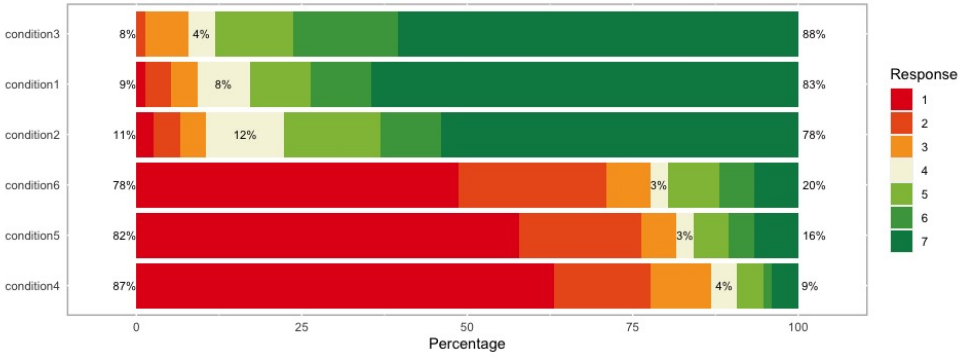


Figure c.4: English-French controls: late bilinguals

d ADVERB PLACEMENT: METHODOLOGICAL COMPARISON

Stadthagen-González et al. (2018) use a different statistical technique to analyse their results than I used in chapter 8. In this section, I analyse my results with the same tests (and their non-parametric equivalents) that were used in Stadthagen-González et al. (2018) (for their AJT data).[1] Stadthagen-González et al. used a one-way ANOVA with post-hoc pairwise comparison. Their pairwise comparison revealed that only the difference between two CS patterns was significant: the equivalents of my conditions 1 and 4.

Note that Stadthagen-González et al. (2018) only tested 4 conditions and did not have equivalents of my conditions 1 and 6. The tests used in this appendix were carried out using the package **rstatix** (Kassambara 2020).

By way of a reminder, the conditions used in the experiments reported in chapter 8 can be found in table d.1.

D.1 DUTCH-ENGLISH

Just like Stadthagen-González et al., the ANOVA found a significant effect of condition (p-value = 6.42e-08). This significant effect was also found with the non-parametric equivalent (the Kruskall-Wallis rank sum test). (Kruskal-Wallis χ^2 = 38.243, df = 5, p-value = 3.371e-07).

1 They analysed their 2AFC data with a different technique. Their results from the 2AFC test showed a significant difference between all conditions except the equivalent of my conditions 4 and 5.

	Dutch-English			English-French		
	verb	adverb	direct object	verb	adverb	direct object
1	Dutch	English	Dutch	French	English	French
2	Dutch	Dutch	English	French	French	English
3	Dutch	English	English	French	English	English
4	English	Dutch	Dutch	English	French	French
5	English	English	Dutch	English	English	French
6	English	Dutch	English	English	French	English

Table d.1: Adverbs in CS: conditions

The results for the t-test pairwise comparisons (parametric) can be found in table d.2 and the results for the Wilcoxon rank sum test (non-parametric) can be found in table d.3. Both were performed with a Bonferroni-correction for multiple comparisons, which is the most conservative correction. As can be seen, the results were the same. Both tests show a statistically significant difference between condition 2 and the others (except for 1). Conditions 5 and 6 are significantly different from 1 and 2 but conditions 3 and 4 are not significant from any others, except 2. So while roughly the same ranking is found as in figure 8.1, it is not statistically confirmed, except for the high ranking of condition 2 and the low ranking of 5 and 6.

Pairwise comparison: Dutch-English

	1	2	3	4	5
2	0.5872	–	–	–	–
3	1	0.0118	–	–	–
4	1	0.0029	1	–	–
5	0.0042	3.1e-07	0.2879	0.7335	–
6	0.0309	5.8e-06	1	1	1

Table d.2: t-test

Pairwise comparison: Dutch-English

	1	2	3	4	5
2	0.76	–	–	–	–
3	1	0.0302	–	–	–
4	1	0.0108	1	–	–
5	0.0055	6.4e-06	0.1990	0.3651	–
6	0.0466	6.0e-05	1	1	1

Table d.3: Wilcoxon rank sum test

D.2 ENGLISH-FRENCH

Both the ANOVA (p-value = 2.29e-06) and Kruskal-Wallis rank sum test showed a significant effect of condition (Kruskal-Wallis χ^2 = 34.36, df = 5, p-value = 2.019e-06). Again, both the post-hoc pairwise comparisons (with Bonferroni corrections) gave the same results (see tables d.4 and d.5). The difference between conditons 1 and 4 & 6 was significant. And just as for the Dutch-English data condition 2 was significantly different from the others (except for 1).

Pairwise comparison: English-French

	1	2	3	4	5
2	1	–	–	–	–
3	0.46042	0.01260	–	–	–
4	0.00656	4.3e-05	1	–	–
5	0.19521	0.00381	1	1	–
6	0.02083	0.00019	1	1	1

Table d.4: t-test

Pairwise comparison: English-French

	1	2	3	4	5
2	1	–	–	–	–
3	0.8138	0.00038	–	–	–
4	0.01399	0.00015	1	–	–
5	0.19405	0.00505	1	1	–
6	0.04269	0.00091	1	1	1

Table d.5: Wilcoxon rank sum test

D.3 SUMMARY

Similarly to what was found by Stadthagen-González et al. (2018), the ANOVA with post-hoc comparison doesn't detect significant differences between many conditions, especially for the Dutch-English data. Stadthagen-González et al. (2018) argue for the use of a different task to better tease apart the differences between these conditions. In chapter 8, I found that statistical modelling also gave an insight into the contribution of the different constituents to the acceptability of sentences containing an AC violation.

BIBLIOGRAPHY

Aaron, Jessi Elana (2015). "Lone English-origin nouns in Spanish: The precedence of community norms." In: *International Journal of Bilingualism* 19.4, pp. 459–480 (cit. on p. 94).

Ackerman, Lauren (2019). "Syntactic and cognitive issues in investigating gendered coreference." In: *Glossa: a journal of general linguistics* 4.1:117 (cit. on pp. 64, 65).

Acquaviva, Paolo (2009). "Roots and lexicality in Distributed Morphology." In: *Fifth York-Essex Morphology Meeting (YEMM), 9th February and 10th February 2008, Department of Language and Linguistic Science, University of York*. University of York. Department of Language and Linguistic Science (cit. on p. 72).

Åfarli, Tor Aanders and Karumuri V Subbarao (2019). "Models of grammar and the outcomes of long-term language contact." In: *Diverse scenarios of syntactic complexity*. Vol. 126. John Benjamins Publishing Company, pp. 27–49 (cit. on pp. 18, 27, 39, 40, 45).

Aguirre, Adalberto Jr (1976). *Acceptability Judgements of Code-Switching Phrases by Chicanos: Some Preliminary Findings*. Stanford University (cit. on pp. 30, 51).

Aikhenvald, Alexandra Y (2016). *How gender shapes the world*. Oxford: Oxford University Press (cit. on p. 141).

Alexiadou, Artemis, Terje Lohndal, Tor Anders Åfarli, and Maren Grimstad Berg (2015). "Language mixing: A Distributed Morphology approach." In: *Proceedings of the 45th annual meeting of the North East Linguistic Society*. Vol. 45. NELS, pp. 25–38 (cit. on p. 98).

Appel, René and Pieter Muysken (1987). *Language Contact and Bilingualism*. London: Edward Arnold (cit. on pp. 27, 29).

Archer, John and Barbara Lloyd (1982). *Sex and gender*. 2nd ed. Cambridge University Press (cit. on p. 64).

Atkinson, Emily (2015). "Gender features on *n* & the root: An account of gender in French." In: *Romance Linguistics 2012: Selected papers from the 42nd Linguistic Symposium on Romance Languages (LSRL), Cedar City, Utah, 20-22 April 2012*. Ed. by Jason Smith and Tabea Ihsane, pp. 229–244 (cit. on pp. 72, 76, 77).

Audring, Jenny (2006). "Pronominal gender in spoken Dutch." In: *Journal of Germanic Linguistics* 18.2, pp. 85–116 (cit. on p. 80).

Audring, Jenny and Geert Booij (2009). "Genus als probleemcategorie." In: *Taal en tongval* 61.1, pp. 13–37 (cit. on p. 80).

Auer, Peter (1988). "A conversation analytic approach to code-switching and transfer." In: Contributions to the Sociology of Language 48. Ed. by Monica Heller, pp. 187–213 (cit. on p. 29).

Auguie, Baptiste (2017). *gridExtra: Miscellaneous Functions for "Grid" Graphics*. R package version 2.3. URL: https://CRAN.R-project.org/package=gridExtra (cit. on p. 115).

Baayen, R H, D J Davidson, and D M Bates (2008). "Mixed-effects modeling with crossed random effects for subjects and items." In: *Journal of Memory and Language* 59, pp. 390–412 (cit. on pp. 56, 57).

Baayen, R.H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge University Press (cit. on pp. 53, 56, 120).

Badiola, Lucia, Rodrigo Delgado, Ariane Sande, and Sara Stefanich (2018). "Code-switching attitudes and their effects on acceptability judgment tasks." In: *Linguistic Approaches to Bilingualism* 8.1, pp. 5–24 (cit. on p. 51).

Badiola, Lucia and Ariane Sande (2018). "Gender assignment in Basque/Spanish mixed determiner phrases: A study of simultaneous bilinguals." In: *Code-switching—Experimental Answers to Theoretical Questions: In Honor of Kay González-Vilbazo*. Ed. by Luis López. Issues in Hispanic and Lusophone Linguistics. Amsterdam: John Benjamins Publishing, pp. 15–61 (cit. on p. 93).

Baker, Mark C (2008). "The macroparameter in a microparametric world." In: *The limits of syntactic variation*. Ed. by Theresa Biberauer. Vol. 132. Linguistics Today. John Benjamins Publishing, pp. 351–374 (cit. on p. 15).

Balam, Osmer (2016). "Semantic categories and gender assignment in contact Spanish: Type of code-switching and its relevance to linguistic outcomes." In: *Journal of Language Contact* 9.3, pp. 405–435 (cit. on pp. 94, 97).

Balam, Osmer, Ukasha Lakshmanan, and Maria del Carmen Parafita Couto (2021). "Gender assignment strategies among simultaneous Spanish/English bilingual children from Miami, Florida." In: *Studies in Hispanic and Lusophone Linguistics* 14 (cit. on p. 93).

Balam, Osmer and Maria del Carmen Parafita Couto (2019). "Adjectives in Spanish/English code-switching: Avoidance of grammatical gender in bi/multilingual speech." In: *Spanish in Context* 16.2, pp. 194–216 (cit. on p. 124).

Bandi-Rao, Shoba and Marcel den Dikken (2014). "Light switches: On *v* as a pivot in codeswitching, and the nature of the ban on word-internal switches." In: *Grammatical theory and bilingual codeswitching*. MIT Press Cambridge, MA, pp. 161–183 (cit. on pp. 43, 44, 51).

Barker, George C (1947). "Social functions of language in a Mexican American community." In: *Acta Americana* 5, pp. 185–202 (cit. on p. 28).

Baron, Naomi S (1971). "A reanalysis of English grammatical gender." In: *Lingua* 27, pp. 113–140 (cit. on p. 80).

Barr, Dale J, Roger Levy, Christoph Scheepers, and Harry J Tily (2013). "Random effects structure for confirmatory hypothesis testing: Keep it maximal." In: *Journal of memory and language* 68.3, pp. 255–278 (cit. on p. 57).

Beatty-Martínez, Anne L, Michelle R Bruni, María Teresa Bajo, and Paola E Dussias (2020). "Brain potentials reveal differential processing of masculine and feminine grammatical gender in native Spanish speakers." In: *Psychophysiology*, e13737 (cit. on p. 96).

Beatty-Martínez, Anne L, Jorge R Valdés Kroff, and Paola E Dussias (2018). "From the field to the lab: a converging methods approach to the study of codeswitching." In: *Languages* 3.2, p. 19 (cit. on p. 46).

Becker, Michael and Michael Dow (under review). "Gender without morphological segmentation in French." In: *Morphology*. https://becker.phonologist.org/projects/frenchgender/becker_dow_french_gender.pdf, consulted 19th of May 2020 (cit. on p. 105).

Belazi, Heidi M, Edward J Rubin, and Almeida Jacqueline Toribio (1994). "Code Switching and X-bar Theory: The Functional Head Constraint." In: *Linguistic Inquiry* 25.2, pp. 221–237 (cit. on pp. 34–36, 164, 165, 170, 172).

Ben-Shachar, Mattan S, Dominique Makowski, and Daniel Lüdecke (2020). "Compute and interpret indices of effect size." In: *CRAN*. R package. URL: https://github.com/easystats/effectsize (cit. on pp. 115, 160).

Benson, Erica J (2001). "The Neglected Early History of Codeswitching Research in the United States." In: *Language & Communication* 21.1, pp. 23–26 (cit. on pp. 27–30, 41).

Bentahila, Abdelâli and Eirlys E Davies (1983). "The Syntax of Arabi-French Code-Switching." In: *Lingua* 59, pp. 301–330 (cit. on pp. 34, 35, 164, 170, 172).

Bermúdez-Otero, Ricardo (2013). "The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features." In: *Probus* 25.1, pp. 3–103 (cit. on p. 70).

Blokzijl, Jeffrey, Margaret Deuchar, and Maria del Carmen Parafita Couto (2017). "Determiner asymmetry in mixed nominal constructions: The role of grammatical factors in data from Miami and Nicaragua." In: *Languages* 2.4, pp. 1–20 (cit. on pp. 34, 97, 101, 105).

Blom, Jan-Petter and John J Gumperz (1972). "Social Meaning in Linguistic Structures: Code Switching in Northern Norway." In: *Directions in Sociolinguistics*. Ed. by John J Gumperz and Dell Hymes. New York: Holt, Rinehart, and Winston, pp. 111–136 (cit. on p. 29).

Bloomfield, Leonard (1935). *Language*. London: George Allen & Unwin LTD (cit. on p. 63).

Bobaljik, Jonathan David (2006). "Where's ϕ? Agreement as a post-syntactic operation." In: *Special Issue of Leiden Papers in Linguistics*. Vol. 20, pp. 1–23 (cit. on p. 72).

Bolonyai, Agnes (2005). "English verbs in Hungarian/English codeswitching." In: *Proceedings of the 4th International Symposium on Bilingualism*. Ed. by Kellie Rolstad and Jeff MacSwan, pp. 317–327 (cit. on pp. 42, 43).

Bolstad, William M and James M Curran (2016). *Introduction to Bayesian statistics*. 3rd ed. John Wiley & Sons (cit. on p. 53).

Boone, Harry N Jr and Deborah A Boone (2012). "Analyzing Likert Data." In: *Journal of Extension* 50.2, 2TOT2 (cit. on p. 55).

Borer, Hagit (2003). "Exo-skeletal vs endo-skeletal explanations: Syntactic projections and the lexicon." In: *The nature of explanation in linguistic theory*. Ed. by John Moore and Maria Polinsky, pp. 31–67 (cit. on p. 18).

Borer, Hagit (2013). *Taking form*. Vol. 3. Structuring sense. Oxford University Press (cit. on p. 18).

Borer, Hagit (2014). "The category of roots." In: *The Syntax of Roots and the Roots of Syntax*. Ed. by Artemis Alexiadou, Hagit Borer, and Florian Schäfer. Oxford University Press (cit. on p. 21).

Box, George EP (1976). "Science and statistics." In: *Journal of the American Statistical Association* 71.356, pp. 791–799 (cit. on p. 56).

Branigan, Holly P. and Martin J. Pickering (2017). "An experimental approach to linguistic representation." In: *Behavioral and Brain Sciences* 40, e282 (cit. on p. 50).

Brasart, Charles (2013). "Structure, variation et configuration du sens dans la parole bilingue Une étude croisée d'alternance codique français-anglais et anglais-allemand." PhD thesis. Université Paris-Sorbonne (cit. on p. 3).

Broekhuis, Hans (2013). *Syntax of Dutch: Adjectives and adjective phrases*. Amsterdam University Press (cit. on pp. 107, 181–183).

Broekhuis, Hans and M E Keizer (2012). *Syntax of Dutch. Nouns and Noun Phrases*. Vol. 1. Amsterdam University Press (cit. on p. 80).

Broersma, Mirjam (2011). "Triggered code-switching: Evidence from picture naming experiments." In: *Modeling bilingualism from structure to chaos: In honor of Kees de Bot*. John Benjamins Amsterdam, pp. 37–57 (cit. on p. 3).

Bross, Fabian (2019a). *Acceptability Ratings in Linguistics: A Practical Guide to Grammaticality Judgments, Data Collection, and Statistical Analysis*. Version 1.02. Mimeo. URL: http://www.fabianbross.de/acceptabilityratings.pdf (cit. on pp. 53, 55, 199).

Bross, Fabian (2019b). *Using Mixed Effect Models to Analyze Acceptability Rating Data in Linguistics*. Version 1.0. Mimeo. URL: http://www.fabianbross.de/mixedmodels.pdf (cit. on pp. 57, 59).

Brown, Becky (1986). "Cajun/English code-switching: A test of formal models." In: *Diversity and diachrony*. Ed. by David Sankoff. Amsterdam: John Benjamins, pp. 399–406 (cit. on p. 3).

Bryer, Jason and Kimberly Speerschneider (2016). *likert: Analysis and Visualization Likert Items*. R package version 1.3.5. URL: https://CRAN.R-project.org/package=likert (cit. on pp. 115, 159, 189).

Burkholder, Michèle (2018). "Language Mixing in the Nominal Phrase: Implications of a Distributed Morphology Perspective." In: *Languages* 3.10, pp. 1–28. DOI: 10.3390/languages3020010. URL: http://www.mdpi.com/2226-471X/3/2/10 (cit. on pp. 69, 97–100, 105, 106, 108).

Burnham, Kenneth P and David R Anderson (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer (cit. on pp. 56–58).

Butt, John, Carmen Benjamin, and Antonia Moreira Rodríguez (2019). *A new reference grammar of modern Spanish*. 6th ed. Routledge (cit. on pp. 68, 69).

Caha, Pavel (2009). "The nanosyntax of case." PhD thesis. Universitetet i Tromsø (cit. on p. 18).

Caha, Pavel (2018). "Notes on insertion in Distributed Morphology and Nanosyntax." In: *Exploring Nanosyntax*. Oxford University Press New York, pp. 57–87. DOI: 10.1093/oso/9780190876746.001.0001 (cit. on p. 18).

Camacho, José and Liliana Sánchez (2017). "Does the verb raise to T in Spanish?" In: *Boundaries, Phases and Interfaces: Case studies in honor of Violeta Demonte*. Ed. by Olga Fernández-Soriano, Elena Castroviejo, and Isabel Pérez-Jiménez. Vol. 239. Linguistics Today. John Benjamins Publishing Company, pp. 47–62 (cit. on p. 183).

Cantone, Katja F (2007). *Code-switching in Bilingual Children*. Dordrecht: Springer (cit. on pp. 153, 165, 170, 176).

Cantone, Katja F and Jeff MacSwan (2009). "Adjectives and word order: A focus on Italian-German." In: *Multidisciplinary Approaches to Code Switching*. Ed. by Ludmila Isurin, David Winford, and Kees de Bot. Amsterdam: John Benjamins Publishing Company, pp. 243–277 (cit. on pp. 38, 51).

Carifio, James and J. Perla Rocco (2007). "Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes." In: *Journal of Social Sciences* 3.3, pp. 106–116 (cit. on p. 54).

Carstens, Vicki (2000). "Concord in minimalist theory." In: *Linguistic inquiry* 31.2, pp. 319–355 (cit. on p. 63).

Carstens, Vicki (2010). "Implications of grammatical gender for the theory of uninterpretable features." In: *Exploring crash-proof grammars*. Ed. by Michael Putnam. Vol. 3. John Benjamins Amsterdam, pp. 31–57 (cit. on p. 71).

Chan, Brian Hok-Shing (2008). "Code-switching, word order and the lexical/functional distinction." In: *Lingua* 118.6, pp. 777–809 (cit. on pp. 153, 166, 170, 176).

Chomsky, Noam (1957/2002). *Syntactic Structures*. 2nd ed. originally published in 1957. Berlin: De Gruyter (cit. on pp. 11, 12).

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press (cit. on pp. 1, 13).

Chomsky, Noam (1970). "Remarks on nominalization." In: *Readings in English Transformational Grammar*. Waltham, pp. 184–221 (cit. on p. 13).

Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris (cit. on pp. 13, 14).

Chomsky, Noam (1995/2015). *The Minimalist Program*. 20th anniversary edition. Cambridge, MA: MIT Press (cit. on p. 15).

Chomsky, Noam (2000a). "Minimalist Inquiries: The Framework." In: *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*. Ed. by Robert Martin, David Michaels, and Juan Uriagereka. Cambridge: MIT Press, pp. 89–155 (cit. on p. 97).

Chomsky, Noam (2000b). "Minimalist inquiries: The framework." In: *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. MIT press, pp. 89–155 (cit. on p. 17).

Chomsky, Noam (2001). "Derivation by Phase." In: *Ken Hale: A life in language*. Ed. by Michael J. Kenstowicz. Current Studies in Linguistics. Cambridge, MA: MIT Press, pp. 1–52 (cit. on pp. 67, 70, 89, 94).

Chomsky, Noam (2008). "On phases." In: *Foundational Issues in Linguistic Theory*. Ed. by Robert Freidin, Carlos P Otero, and Maria Luisa Zubizarreta. Vol. 45. Current Studies in Linguistics. MIT press, pp. 133–166 (cit. on pp. 17, 175).

Christensen, R. H. B. (2019). *ordinal—Regression Models for Ordinal Data*. R package version 2019.12-10. https://CRAN.R-project.org/package=ordinal (cit. on pp. 160, 189).

Citko, Barbara (2014). *Phase theory: An introduction*. Cambridge University Press (cit. on p. 17).

Claes, Jeroen (2020). *Statitics for Linguistics*. consulted June 2020. URL: http://www.jeroenclaes.be/statistics_for_linguistics/ (cit. on pp. 53, 58).

Cohen, Jacob (1988). *Statistical power analysis for the behavioral sciences*. Academic press (cit. on pp. 117, 162, 191).

Comrie, Bernard, Martin Haspelmath, and Balthasar Bickel (2008). *The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses*. Retrieved 02-09-2020. Department of Linguistics of the Max Planck Institute for Evolutionary Anthropology & the Department of Linguistics of the University of Leipzig. URL: https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf (cit. on p. xxiii).

Conrod, Kirby (2019). "Pronouns raising and emerging." PhD thesis. University of Washington (cit. on p. 65).

Conrod, Kirby (2020). *nonbinary_names*. Retrieved 17-09-2020. URL: https://docs.google.com/spreadsheets/d/1GF6c5qFFzTqYGukRYia8WcSam48tBHm_R6MJB5tJPiI/edit#gid=0 (cit. on p. 53).

Corbett, Greville G. (1991). *Gender*. Cambridge Textbooks in Linguistics. Cambridge, NY: Cambridge University Press (cit. on pp. 61, 63, 64, 66, 74, 75, 141).

Corbett, Greville G and Norman M Fraser (1999). "Default genders." In: *Gender in Grammar and Cognition*. Ed. by Barbara Unterbeck and Matti Rissanen. Vol. 21. Trends in Linguistics: Studies and Monographs, pp. 55–97 (cit. on p. 96).

Cornips, Leonie and Gunther De Vogelaer (2009). "Variatie en verandering in het Nederlandse genus: een multidisciplinair perspectief." In: *Taal en tongval* 61.1, pp. 1–12 (cit. on p. 81).

Cowart, Wayne (1997). *Experimental syntax*. Sage (cit. on p. 50).

Cunnings, Ian (2012). "An overview of mixed-effects statistical models for second language researchers." In: *Second Language Research* 28.3, pp. 369–382 (cit. on pp. 56, 58).

De Belder, Marijke (2011). "Roots and Affixes: Eliminating Categories from Syntax." PhD thesis. Utrecht University (cit. on p. 21).

De Belder, Marijke and Jeroen Van Craenenbroek (2015). "How to merge a root." In: *Linguistic Inquiry* 46.4, pp. 625–655 (cit. on pp. 21, 72).

De Caluwe, Johan (2009). "Tussentaal wordt omgangstaal in Vlaanderen." In: *Nederlandse taalkunde* 14.1, pp. 8–25 (cit. on p. 4).

De Schepper, Philippe, Pieter De Graeve, Mathias Claeys, Bas Adriaensen, and Joël Vanhoebrouck (2019). "No kotje kotje (Television series episode)." In: *Eigen kweek*. Eyeworks Film & TV Drama (produced by Katrien Carton) (cit. on p. 84).

De Schutter, Georges (2005). *Morfologische atlas van de Nederlandse Dialecten: Deel 1: meervoudsvorming bij zelfstandig naamwoorden, vorming van verkleinwoorden, geslacht bij zelfstandig naamwoord, bijvoeglijk naamwoord en bezittelijk Voornaamwoord*. Vol. 1. Amsterdam University Press (cit. on p. 82).

De Vogelaer, Gunther (2010). "(Not) acquiring grammatical gender in two varieties of Dutch." In: *Advances in cognitive sociolinguistics*. Ed. by Dirk Geeraerts, Gitte Kristiansen, and Yves Peirsman. Mouton de Gruyter Berlin, pp. 167–190 (cit. on p. 80).

De Vogelaer, Gunther and Gert De Sutter (2011). "The geography of gender change: Pronominal and adnominal gender in Flemish dialects of Dutch." In: *Language Sciences* 33.1, pp. 192–205 (cit. on p. 80).

De Winter, Joost C F and Dimitra Dodou (2010). "Five-point Likert items: t-test versus Mann-Whitney-Wilcoxon." In: *Practical Assessment, Research & Evaluation* (cit. on p. 55).

Delbeke, Tijs and Alban Sarens (2011). *Waar is de meisje [recorded by De Hoop]*. EMI (cit. on p. 125).

Delgado, Rodrigo (2018). "The familiar and the strange: Gender assignment in Spanish/Englis mixed DPs." In: *Code-switching – Experimental Answers to Theoretical Questions: In honor of Kay González-Vilbazo*. Ed. by Luis López. Vol. 19. Issues in Hispanic and Lusophone Linguistics. Amsterdam: John Benjamins Publishing Company, pp. 39–62 (cit. on pp. 94, 97, 104).

Den Dikken, Marcel (2013a). "Introduction." In: *The Cambridge handbook of Generative syntax*. Ed. by Marcel Den Dikken. Cambridge University Press (cit. on p. 12).

Den Dikken, Marcel (2013b). "The Cambridge handbook of Generative syntax." In: Cambridge University Press (cit. on p. 11).

Desagulier, Guillaume (2017). *Corpus Linguistics and Statistics with R*. Springer (cit. on p. 53).

Deuchar, Margaret (2006). "Welsh-English code-switching and the Matrix Language Frame model." In: *Lingua* 116.11, pp. 1986–2011 (cit. on p. 97).

Deuchar, Margaret (2020). "Code-Switching in Linguistics: A Position Paper." In: *Languages* 5.2, pp. 1–22 (cit. on pp. 34, 42, 43).

Di Sciullo, Anne-Marie, Pieter Muysken, and Rajendra Singh (Mar. 1986). "Government and Code-Mixing." In: *Journal of Linguistics* 22.1, pp. 1–24 (cit. on pp. 34–36).

Diercks, Michael (2011). "Against Feature Inheritance: Phase heads are not defined by unvalued phi features." unpublished manuscript. Claremont, CA (cit. on p. 175).

Don, Jan (2005). "Roots, deverbal nouns and denominal verbs." In: *Morphology and Linguistic Typology, On-line Proceedings of the Fourth Mediterranean Morphology Meeting (MMM4)*. Ed. by Geert Booij, E Guevara, A Rally, and S Scalise. Vol. 4, pp. 91–104 (cit. on pp. 24, 86).

Ebert, Shane and Bradley Hoot (2018). "*That*-trace effects in Spanish-English code-switching." In: *Code-switching–Experimental Answers to Theoretical Questions: In honor of Kay González-Vilbazo*. Ed. by Luis López. Vol. 19. Issues in Hispanic and Lusophone Linguistics. John Benjamins Publishing Company, p. 101 (cit. on pp. 165, 175, 176).

Ebert, Shane and Bryan Koronkiewicz (2018). "Monolingual stimuli as a foundation for analyzing code-switching data." In: *Linguistic Approaches to Bilingualism* 8.1, pp. 25–66 (cit. on pp. 52, 114, 187).

Embick, David (2015). *The morpheme: A theoretical introduction*. Vol. 31. Interface Explorations. Walter de Gruyter GmbH & Co KG (cit. on pp. 11, 18).

Endresen, Anna and Laura Alexis Janda (2016). "Five statistical models for Likert-type experimental data on acceptability judgments." In: *Journal of Research Design and Statistics in Linguistics and Communication Science* (2), pp. 217–250 (cit. on p. 55).

Espinosa, Aurelio M (1914). "Studies in New Mexican Spanish, Part III: The English elements (Part of doctoral dissertation, University of Chicago, 1909)." In: *Revue de Dialectologie Romane* 6, pp. 241–317 (cit. on pp. 28, 30, 41).

Fabozzi, Frank J, Sergio M Focardi, Svetlozar T Rachev, and Bala G Arshanapalli (2014). *The basics of financial econometrics: Tools, concepts, and asset management applications*. John Wiley & Sons (cit. on p. 58).

Fathi, Radwa and Jean Lowenstamm (2016). "The gender assignment pattern of French nouns." In: *Morphology* 26.3-4, pp. 477–509 (cit. on pp. 77–79, 88).

Ferguson, Christopher J (2009). "An effect size primer: A guide for clinicians and researchers." In: *Professional Psychology: Research and Practice* (cit. on p. 59).

Fillmore, Charles J (1992). ""Corpus Linguistics" or "Computer-aided armchair linguistics"." In: *Directions in corpus linguistics. Proceedings of Nobel Symposium*. Vol. 82, pp. 35–60 (cit. on p. 205).

Flom, George T (1903). "The gender of English loan-nouns in Norse dialects in America; A contribution to the study of the development of grammatical gender." In: *The Journal of English and Germanic Philology* 5.1, pp. 1–31 (cit. on p. 141).

Flores, Abel Cruz (2018). "Gender Features in Spanish-English Asymmetrical Codeswitches" (cit. on p. 97).

Fraser, Norman M. and Greville G. Corbett (1997). "Defaults in Arapesh." In: *Lingua* 103.1, pp. 25–57 (cit. on p. 109).

Gallego, Ángel J (2014). "Deriving feature inheritance from the copy theory of movement." In: *The Linguistic Review* 31.1, pp. 41–71 (cit. on pp. 175, 176, 203).

Ganushchak, Lesya Y, Rinus G Verdonschot, and Niels O Schiller (2011). "When leaf becomes neuter: event-related potential evidence for grammatical gender transfer in bilingualism." In: *Neuroreport* 22.3, pp. 106–110 (cit. on p. 88).

Gardner-Chloros, Penelope (2009). *Code-Switching*. New York: Cambridge University Press (cit. on pp. 25, 28, 29).

Gibson, Edward, Steve Piantadosi, and Kristina Fedorenko (2011). "Using Mechanical Turk to obtain and analyze English acceptability judgments." In: *Language and Linguistics Compass* 5.8, pp. 509–524 (cit. on p. 59).

Gingràs, Rosario (1974). "Problems in the description of Spanish-English intrasentential code-switching." In: *Southwest areal linguistics*, pp. 167–174 (cit. on p. 30).

González-Vilbazo, Kay (2005). "Die Syntax des Code-Switching." PhD thesis. University of Cologne (cit. on p. 165).

González-Vilbazo, Kay and Bryan Koronkiewicz (2016). "*Tú y yo* can codeswitch, *nosotros* cannot." In: *Spanish English Codeswitching Caribbean US*. Vol. 11, pp. 237–260 (cit. on p. 31).

González-Vilbazo, Kay and Luis López (2012). "Little *v* and parametric variation." In: *Natural Language and Linguistic Theory* 30.1, pp. 33–77 (cit. on pp. 38, 167).

Goto, Nobu (2011). "Some Consequences of Feature Inheritance." In: *Studies in English Literature: Regional Branches Combined Issue*. Ed. by The Tohoku branch of the English Literary Society of Japan. Vol. 3, pp. 113–136 (cit. on p. 175).

Grabowski, Jane (May 2011). "Stability of Grammaticality Judgments in German-English Code-Switching." PhD thesis. Arizona State University (cit. on p. 51).

Gries, Stefan Th. (2009). *Statistics for linguistics with R: A practical introduction*. Walter de Gruyter (cit. on pp. 53, 56).

Grimstad, Maren Berg (2017). "The code-switching/borrowing debate: evidence from English-origin verbs in American Norwegian." In: *Lingue e linguaggio* 16.1, pp. 3–34 (cit. on pp. 41, 43, 44).

Grimstad, Maren Berg, Terje Lohndal, and Tor Anders Åfarli (2014). "Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian." In: *Nordlyd* 41.2, pp. 213–237 (cit. on pp. 27, 39, 98).

Grimstad, Maren Berg, Brita Ramsevik Riksem, Terje Lohndal, and Tor Anders Åfarli (2018). "Lexicalist vs exoskeletal approaches to lan-

guage mixing." In: *The Linguistic Review* 35.2, pp. 187–218 (cit. on p. 98).

Grondelaers, Stefan and Roeland van Hout (2011). "The standard language situation in the Low Countries: Top-down and bottom-up variations on a diaglossic theme." In: *Journal of Germanic Linguistics* 23 (3), pp. 199–243. DOI: 10.1017/S1470542711000110 (cit. on p. 4).

Gross, Steven (in press). "Linguistic Judgments as Evidence." In: *The Blackwell Companion to Chomsky*. Ed. by Nicholas Allott, Lohndal Terje, and Georges Rey. Wiley-Blackwell (cit. on p. 50).

Gullberg, Marianne, Peter Indefrey, and Pieter Muysken (2009). "Research techniques for the study of code-switching." In: *The Cambridge handbook on lingusitic code-switching*. Ed. by Barbara E Bullock and Almeida Jacqueline Toribio. Cambridge: Cambridge University Press, pp. 21–39 (cit. on p. 46).

Haegeman, Liliane (1994). *Introduction to government and binding theory*. 2nd ed. Wiley-Blackwell (cit. on p. 13).

Haegeman, Liliane and Marjo Van Koppen (2012). "Complementizer agreement and the relation between C^0 and T^0." In: *Linguistic Inquiry* 43.3, pp. 441–454 (cit. on p. 175).

Halle, Morris and Alec Marantz (1993). "Distributed Morphology and the Pieces of Inflection." In: *The view from building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Ed. by Kenneth Locke Hale and Samuel Jay Keyser. Cambridge, MA: MIT Press, pp. 111–176 (cit. on pp. 18, 38).

Halle, Morris and Alec Marantz (1994). "Some key features of Distributed Morphology." In: *MIT working papers in linguistics* 21, pp. 275–288 (cit. on pp. 18, 22).

Harley, Heidi (2014). "On the identity of roots." In: *Theoretical Linguistics* 40.3/4, pp. 225–276 (cit. on p. 21).

Harley, Heidi and Rolf Noyer (Apr. 1999). "Distributed Morphology." In: *Glot International* 4.4, pp. 3–9 (cit. on pp. 11, 18, 20, 23).

Harris, James W. (1991). "The Exponence of Gender in Spanish." In: *Linguistic Inquiry* 22.1, pp. 27–62 (cit. on pp. 92, 96).

Harris, James W (1996). "The syntax and morphology of class marker suppression in Spanish." In: *Grammatical theory and Romance languages: selected papers from the 25th linguistic symposium on Romance languages (LSRL XXV)*. Ed. by Karen Zagona. Amsterdam: John Benjamins, pp. 99–122 (cit. on pp. 18, 69, 72).

Haugen, Einar I (1953a). *Bilingualism in the Americas: A Bibliography and Research Guide*. Vol. 26. Publication of the American Dialect Society (cit. on p. 29).

Haugen, Einar I (1953b). *The Norwegian Language in America: : A Study in Bilingual Behavior*. Publications of the American Institute, University of Oslo. University of Pennsylvania Press (cit. on p. 29).

Herring, Jon Russel, Margaret Deuchar, Maria del Carmen Parafita Couto, and Monica Moro Quintanilla (Aug. 2010). "'I saw the *madre*': evaluating predictions about codeswitched determiner-noun sequences using Spanish-English and Welsh-English data." In: *International Journal of Bilingual Education and Bilingualism* 12.5, pp. 553–573 (cit. on pp. 34, 38).

Heumann, Christian, Michael Schomaker, and Shalabh (2016). *Introduction to statistics and data analysis*. Springer (cit. on pp. 53, 54).

Hill, Joseph C, Diane C Lillo-Martin, and Sandra K Wood (2018). *Sign Languages: Structures and Contexts*. Routledge Guides to Linguistics. Routledge (cit. on p. 27).

Hjelmslev, Louis (1956). "Animé et inanimé, personnel et non-personnel." In: *Essais linguistiques*. Vol. 12. Travaux de l'Institut de Linguistique, pp. 211–249 (cit. on p. 141).

Holmberg, Anders (2015). "Verb Second." In: *Syntax – Theory and Analysis. An International Handbook*. Ed. by T Kiss and Artemis Alexiadou. Vol. 42. Handbooks of Linguistics and Communication Science. Berlin: De Gruyter, pp. 342–382 (cit. on pp. 146, 147, 152).

Hoot, Brad (2011). *Complementizer asymmetry in Spanish-English codeswitching*. Oral presentation at International Symposium on Bilingualism (ISB) 8 (cit. on p. 165).

Ibrahim, Muhammad Hassan (1973). *Grammatical gender: Its origin and development*. Berlin (cit. on p. 63).

Itô, Junko and Armin Mester (2003). "On the sources of opacity in OT: Coda processes in German." In: *The syllable in optimality theory*, pp. 271–303 (cit. on p. 43).

Jake, Janice L, Carol Myers-Scotton, and Steven Gross (2002). "Making a minimalist approach to codeswitching work: Adding the Matrix Language." In: *Bilingualism: Language and Cognition* 5, pp. 69–91 (cit. on pp. 34, 43, 93).

Jake, Janice L, Carol Myers-Scotton, and Steven Gross (2005). "A response to MacSwan (2005): Keeping the Matrix Language." In: *Bilin-*

gualism: Language and Cognition 8.3, pp. 271–276 (cit. on pp. 42, 43, 97).

James, Allan (2016). "From code-mixing to mode-mixing in the European context." In: *World Englishes* 35.2, pp. 259–275 (cit. on p. 29).

Jansen, Veronica, Jasmin Müller, and Natascha Müller (2012). "Code-switching between en OV and VO language: Evidence from German-Italian, German-French and German-Spanish children." In: *Linguistic Approaches to Bilingualism* 2.4, pp. 337–378 (cit. on pp. 150, 153, 165, 170, 176).

Jaspers, Jürgen and Sarah Van Hoof (2015). "Ceci n'est pas une Tussentaal: evoking standard and vernacular language through mixed Dutch in Flemish telecinematic discourse." In: *Journal of Germanic Linguistics* 27.1, pp. 1–44 (cit. on pp. 4, 5).

Jiménez-Fernández, Ángel L and Shigeru Miyagawa (2014). "A feature-inheritance approach to root phenomena and parametric variation." In: *Lingua* 145, pp. 276–302 (cit. on p. 175).

Jorschick, Liane, Antje Endesfelder Quick, Dana Glässer, Elena Lieven, and Michael Tomasello (2011). "German–English-speaking children's mixed NPs with 'correct' agreement." In: *Bilingualism: Language and Cognition* 14.2, pp. 173–183 (cit. on p. 97).

Joshi, Aravind K (1982). "Processing of Sentences with Intra-Sentential Code-Switching." In: *Proceedings of the 9th conference on Computational linguistics*. Ed. by J. Horecký, pp. 145–150 (cit. on pp. 34–36).

Joshi, Aravind K (1985). "Processing of sentences with intra-sentential code-switching." In: *Natural language parsing: Psycholinguistic, computational and theoretical perspectives*. Cambridge: Cambridge University Press, pp. 190–200 (cit. on p. 32).

Joshi, Aravind K (1987). "An introduction to tree adjoining grammars." In: *Mathematics of language* 1, pp. 87–115 (cit. on p. 36).

Judd, Charles M, Jacob Westfall, and David A Kenny (2017). "Experiments with more than one random factor: Designs, analytic models, and statistical power." In: *Annual Review of Psychology* 68, pp. 601–625 (cit. on pp. 121, 163, 191).

Karmiloff-Smith, Annette (1981). *A functional approach to child language: A study of determiners and reference*. Vol. 24. Cambridge, UK: Cambridge University Press (cit. on p. 75).

Kassambara, Alboukadel (2020). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests*. R package version 0.5.0. URL: https://CRAN.R-project.org/package=rstatix (cit. on p. 227).

Kayne, Richard S (1994). *The Antisymmetry of Syntax*. Vol. 25. Linguistic Inquiry Monographs. Cambridge, MA: The MIT Press (cit. on p. 153).

Kinyon, Alexandra, Tatjana Scheffler, Owen Rambow, Yoon SinWon, and Aravind K Joshi (2006). "The Metagrammar Goes Multilingual: A Cross-Linguistic Look at the V2-Phenomenon." In: *Proceedings of the 8th International Workshop on Tree Adjoining Grammar and Related Formalisms*, pp. 17–24 (cit. on p. 171).

Kizach, Johannes (2014). *Analyzing Likert-scale data with mixed-effects linear models: a simulation study*. Poster Presented at Linguistic Evidence 2014. URL: http://pure.au.dk/portal/files/70360382/SimulationPosterJK.pdf (cit. on p. 59).

Klassen, Rachel and Juana M Liceras (2017). "The representation of gender in the mind of Spanish-English bilinguals: Insights from code-switched Adjectival Predicates." In: *Borealis – An International Journal of Hispanic Linguistics* 6.1, pp. 77–96 (cit. on pp. 63, 93).

Kootstra, Gerrit Jan, Janet G van Hell, and Ton Dijkstra (2010). "Syntactic alignment and shared word order in code-switched sentence production: Evidence from bilinugal monologue and dialogue." In: *Journal of Memory and Language* 63, pp. 210–231 (cit. on pp. 151, 153).

Koronkiewicz, Bryan (2019a). "Control Stimuli in Experimental Code-Switching Research." In: *Languages* 4.3, p. 67 (cit. on pp. 51, 52, 113, 131, 157, 186, 187, 193, 198, 209, 223–225).

Koronkiewicz, Bryan (2019b). *Testing adverb position in in Spanish-English code-switching: Evidence from an acceptability judgment task*. Oral presentation at the Conference on Multilingualism (COM) 2019, Leiden University (cit. on pp. 183, 184, 194).

Koronkiewicz, Bryan and Shane Ebert (2018). "Modality in experimental code-switching research." In: *Code-switching—Experimental Answers to Theoretical Questions: In Honor of Kay González-Vilbazo*. Ed. by Luis López. Issues in Hispanic and Lusophone Linguistics. Amsterdam: John Benjamins Publishing, pp. 147–76 (cit. on p. 51).

Koster, Jan (1975). "Dutch as a SOV Language." In: *Linguistic analysis* 1, pp. 111–136 (cit. on pp. 145, 181).

Kotek, Hadas, Sarah Babinski, Rikker Dockum, and Christopher Geissler (2020). "Gender representation in linguistic example sentences." In: *Proceedings of the Linguistic Society of America*. Vol. 5. 1, pp. 514–528 (cit. on p. 53).

Kramer, Ruth (2009). "Definite markers, phi-features, and agreement: A morphosyntactic investigation of the Amharic DP." PhD thesis. Santa Cruz, CA: University of California (cit. on pp. 72, 77, 88).

Kramer, Ruth (2015). *The Morphosyntax of Gender*. eng. Oxford UP. ISBN: 9780199679942 (cit. on pp. 63, 67–74, 84, 85, 88, 96, 98–100, 108, 110, 128, 201).

Kramer, Ruth (2016). "The location of gender features in the syntax." In: *Language and linguistics Compass* 10.11, pp. 661–677 (cit. on pp. 72, 79).

Labov, William (1971). "The notion of 'system' in creole studies." In: *Pidginization and creolization of languages*. Ed. by Dell Hymes. Vol. 447. Cambridge University Press Cambridge, p. 472 (cit. on p. 30).

Labov, William (1972). *Sociolinguistic patterns*. Philadelphia, PA: University of Pennsylvania Press (cit. on p. 50).

Laenzlinger, Christopher (1998). *Comparative studies in word order variation: Adverbs, pronouns, and clause structure in Romance and Germanic*. Vol. 20. John Benjamins Publishing (cit. on p. 181).

Lasnik, Howard and Terje Lohndal (2013). "Brief overview of the history of generative syntax." In: *The Cambridge handbook of Generative syntax*. Ed. by Marcel Den Dikken. Cambridge University Press (cit. on p. 12).

Lee, Tong-King (2015). "Translanguaging and visuality: Translingual practices in literary art." In: *Applied Linguistics Review* 6.4, pp. 441 –465. DOI: https://doi.org/10.1515/applirev-2015-0022. URL: https://www.degruyter.com/view/journals/alr/6/4/article-p441.xml (cit. on p. 29).

Lehtinen, Meri (1966). "An analysis of a Finnish-English bilingual corpus." PhD thesis. University of Indiana (cit. on pp. 30, 31).

Lenth, Russell (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.4.7. URL: https://CRAN.R-project.org/package=emmeans (cit. on pp. 121, 197).

Liceras, Juana M, Raquel Fernández Fuertes, and Rachel Klassen (2016). "Language dominance and language nativeness." In: *Spanish-English code-switching in the Caribbean and the US*. Vol. 11. Issues in Hispanic and Lusophone Linguistics. John Benjamins Publishing Company, pp. 107–138 (cit. on p. 93).

Liceras, Juana M, Raquel Fernández Fuertes, Susana Perales, Rocío Pérez-Tattam, and Kenton Todd Spradlin (2008). "Gender and gender agreement in bilingual native and non-native grammars: A view from child

and adult functional-lexical mixings." In: *Lingua* 118, pp. 827–851 (cit. on pp. 39, 93, 97–99, 104, 105).

Likert, Rensis (1932). "A technique for the measurement of attitudes." In: *Archives of psychology* (cit. on p. 54).

LimeSurvey Development Team (2012). *LimeSurvey - The free and open source survey software tool!* LimeSurvey Development Team. URL: limesurvey.org (cit. on pp. 114, 188, 194).

Lochbihler, Bethany and Eric Mathieu (2016). "Clause typing and feature inheritance of discourse features." In: *Syntax* 19.4, pp. 354–391. DOI: 10.1111/synt.12126 (cit. on p. 175).

Lohndal, Terje (2013). "Generative grammar and language mixing." In: *Theoretical Linguistics* 39.3-4, pp. 215–224 (cit. on p. 1).

López, Luis (2007). *Locality and the architecture of syntactic dependencies*. Pagrave MacMillan (cit. on p. 103).

López, Luis (Mar. 2014). "Updates from the UIC Center for Bilingualism." Oral presentation at the LACG-talks, Leiden University (cit. on pp. 115, 156).

López, Luis (2018). *Code-switching – Experimental Answers to Theoretical Questions: In honor of Kay González-Vilbazo*. Vol. 19. Issues in Hispanic and Lusophone Linguistics. John Benjamins Publishing Company (cit. on p. 51).

López, Luis (2020). *Bilingual grammar: Toward an integrated model*. Cambridge University Press (cit. on pp. 39, 40, 49, 97, 103, 175).

López, Luis, Artemis Alexiadou, and Tonjes Veenstra (2017). "Code-switching by phase." In: *Languages* 2.3, p. 9 (cit. on pp. 17, 40, 44, 45).

Loporcaro, Michele and Tania Paciaroni (2011). "Four-gender systems in Indo-European." In: *Folia linguistica* 45.2, pp. 389–433 (cit. on p. 79).

Lowenstamm, Jean (2012). "Feminine and gender, or why the feminine profile of French nouns has nothing to do with gender." In: *Linguistic inspirations. Edmund Gussmann in memoriam*. Wydawnictwo Katolicki Uniwersytet Lubelski, Lublin, pp. 371–406 (cit. on pp. 21, 24).

Luque, Alicia, Nethaum Mizyed, and Kara Morgan-Short (2018). "Event-related potentials reveal evidence for syntactic co-activation in bilingual language processing." In: *Code-switching–Experimental Answers to Theoretical Questions: In honor of Kay González-Vilbazo*. Ed. by Luis López. Vol. 19. John Benjamins Publishing Company, pp. 177–194 (cit. on p. 46).

Luraghi, Silvia and Claudia Parodi (2008). *Key terms in syntax and syntactic theory*. Bloomsbury Publishing (cit. on pp. 11, 17, 67).

Lyster, Roy (2006). "Predictability in French gender attribution: A corpus analysis." In: *Journal of French Language Studies* 16.1, pp. 69–92 (cit. on p. 75).

MacSwan, Jeff (1999). *A minimalist approach to intrasentential code switching*. Outstanding Dissertations in Linguistics. New York: Garland Publishing, Inc (cit. on pp. 37, 39, 40, 42).

MacSwan, Jeff (2004). "Code switching and linguistic theory." In: *Handbook of Bilingualism*. Ed. by Tej K Bhatia and William C Ritchie. Oxford: Blackwell. Chap. 11, pp. 415–462 (cit. on pp. 153, 170, 176).

MacSwan, Jeff (2005). "Codeswitching and generative grammar: A critique of the MLF model and some remarks on "modified minimalism"." In: *Bilingualism: Language and Cognition* 8, pp. 1–22 (cit. on pp. 99, 105).

MacSwan, Jeff (2009). "Generative approaches to code-switching." In: *The Cambridge Handbook of Lingusitic Code-switching*. Ed. by Barbara E Bullock and Almeida Jacqueline Toribio. Cambridge: Cambridge University Press, pp. 209–335 (cit. on pp. 37, 38).

MacSwan, Jeff (2014). "Programs and Proposals in Codeswitching Research: Unconstraining Theories of Bilingual Language Mixing." In: *Grammatical Theory and Bilingual Codeswitching*. Ed. by Jeff MacSwan. Cambridge, MA: MIT Press, pp. 1–34 (cit. on p. 172).

MacSwan, Jeff and Sonia Colina (2014). "Some consequences of language design: Code switching and the PF interface." In: *Grammatical Theory and Bilingual Codeswitching*. Ed. by Jeff MacSwan. Cambridge, MA, USA: MIT Press, pp. 185–200 (cit. on pp. 42, 171, 172).

Macaulay, Monica and Colleen Brice (1997). "Don't touch my projectile: Gender bias and stereotyping in syntactic examples." In: *Language*, pp. 798–825 (cit. on p. 52).

Mahootian, Shahrzad (1993). "A null theory of code switching." PhD thesis. Northwestern University. (cit. on pp. 36, 37, 153, 170, 176).

Mahootian, Shahrzad (2005). "Linguistic change and social meaning: Codeswitching in the media." In: *International Journal of Bilingualism* 9, pp. 361–375 (cit. on p. 51).

Mahootian, Shahrzad and Beatrice Santorini (1996). "Code Switching and the Complement/Adjunct Distinction." In: *Linguistic Inquiry* 27.3, pp. 464–479 (cit. on pp. 153, 164, 166, 170, 172, 176).

Maiden, Martin (2011). "Morphological persistence." In: *The Cambridge history of the Romance languages*. Ed. by Martin Maiden, John Charles

Smith, and Adam Ledgeway. Vol. 1. Cambridge University Press Cambridge, pp. 155–215 (cit. on p. 74).

Matasović, Ranko (2004). *Gender in Indo-European*. Universitaetsverlag Winter (cit. on p. 141).

Matuschek, Hannes, Reinhold Kliegl, Shravan Vasishth, Harald Baayen, and Douglas Bates (2017). "Balancing Type I error and power in linear mixed models." In: *Journal of Memory and Language* 94, pp. 305–315 (cit. on p. 57).

Mazerolle, Marc J (2019). *AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c)*. R package version 2.2-2. URL: https://cran.r-project.org/package=AICcmodavg (cit. on p. 189).

McClure, Erica F (1977). "Aspects of code-switching in the discourse of bilingual Mexican-American children." In: ed. by Muriel Saville-Troike. Georgetown University Press, pp. 93–115 (cit. on p. 30).

Meisel, Jürgen M (2004). "The bilingual child." In: *The handbook of bilingualism*, pp. 91–113 (cit. on p. 149).

Mendenhall, William, Robert J Beaver, and Barbara M Beaver (2012). *Introduction to probability and statistics*. Cengage Learning (cit. on p. 53).

Meyerhoff, Miriam (2018). *Introducing sociolinguistics*. 3rd ed. Routledge (cit. on pp. 26, 27, 29).

Montes-Alcalá, Cecilla (2000). "Written codeswitching. Powerful bilingual images." In: *Codeswitching Worldwilde II*. Ed. by Rodolfo Jacobson. Trends in Linguistics. Studies and Monographs 126. Berlin: Mouton de Gruyter, pp. 193–219 (cit. on p. 51).

Müller, Gereon (2004). "Verb-Second as vP-First." In: *Journal of Comparative Germanic Linguistics* 7, pp. 179–234 (cit. on p. 149).

Muysken, Pieter (1997). "Code-switching processes: Alternation, insertion, congruent lexicalization." In: *Language choices: Conditions, constraints, and consequences*. Vol. 1. Studies in Language and Society. John Benjamins Amsterdam, pp. 361–380 (cit. on p. 26).

Muysken, Pieter (2000). *Bilingual Speech: A typology of code-mixing*. Cambridge, UK: Cambridge University Press (cit. on p. 44).

Myers-Scotton, Carol M and Janice L Jake (2017). "Revisiting the 4-M model: Codeswitching and morpheme election at the abstract level." In: *International Journal of Bilingualism* 21.3, pp. 340–366 (cit. on pp. 33, 34).

Myers-Scotton, Carol (1993/1997). *Duelling languages: Grammatical structure in codeswitching*. paperback edition. Oxford: Oxford University Press (cit. on pp. 32, 34).

Myers-Scotton, Carol (2006). *Multiple Voices: an introduction to bilingualism*. Cornwall: Blackwell Publishing Ltd (cit. on pp. 32, 33).

Myers-Scotton, Carol and Janice Jake (2009). "A universal model of code-switching and bilingual language processing of production." In: *The Cambridge Handbook of Lingusitic Code-switching*. Ed. by Barbara E Bullock and Almeida Jacqueline Toribio. New York: Cambridge University Press, pp. 336–357 (cit. on pp. 33, 34).

Nieminen, Pentti, Heli Lehtiniemi, Kirsi Vähäkangas, Antti Huusko, and Arja Rautio (2013). "Standardised regression coefficient as an effect size index in summarising findings in epidemiological studies." In: *Epidemiology, Biostatistics and Public Health* 10.4, pp. 1–15 (cit. on p. 59).

Norman, Geoff (2010). "Likert scales, levels of measurement and the "laws" of statistics." In: *Advances in health sciences education* 15.5, pp. 625–632 (cit. on p. 55).

Norris, Mark (2014). "A theory of nominal concord." PhD thesis. UC Santa Cruz (cit. on p. 63).

Pablos, Leticia et al. (2019). "Adjective-noun order in Papiamento-Dutch code-switching." In: *Linguistic Approaches to Bilingualism* 9.4–5, pp. 710–735 (cit. on p. 38).

Parafita Couto, Maria del Carmen (2019). *Code-switching across communities and language pairs*. Keynote address at the Conference on Multilingualism (COM) 2019, Leiden University (cit. on pp. 92, 97).

Parafita Couto, Maria del Carmen, Amaia Munarriz, Irantzu Epelde, Margaret Deuchar, and Beñat Oyharçabal (2015). "Gender conflict resolution in Spanish-Basque mixed DPs." In: *Bilingualism: Language and Cognition* 18.2, pp. 304–323. DOI: 10.1017/S136672891400011X (cit. on p. 93).

Parafita Couto, Maria del Carmen and Eva Rodríguez-González (2019). "Do all bilinguals switch the same way? Testing patterns of non-habitual Spanish-English code-switchers." In: *Espacios en evolución: confluencias lingüísticas y culturales*. Ed. by Efthimia Pandis Pavlakis, Haralambos Symeonidis, Paul Chandler, María Tsokou, and Viktoria Kritikou. Madrid: Ediciones del Orto, pp. 105–127 (cit. on p. 51).

Pennycook, Alastair (2016). "Mobile times, mobile terms: The trans-super-poly-metro movement." In: *Sociolinguistics: theoretical de-*

bates. Cambridge University Press Cambridge, pp. 201–216 (cit. on p. 27).

Pérez-Pereira, Miguel (1991). "The acquisition of gender: What Spanish children tell us." In: *Journal of Child Language* 18.3, pp. 571–590 (cit. on p. 75).

Pesetsky, David and Esther Torrego (2007). "The syntax of valuation and the interpretability of features." In: *Phrasal and clausal architecture: Syntactic derivation and interpretation*. Benjamins, pp. 262–294 (cit. on p. 71).

Pfaff, Carol W (June 1979). "Constraints on Language Mixing: Intrasentential Code-Switching and Borrowing in Spanish/English." In: *Language* 55.2, pp. 291–318 (cit. on pp. 30, 31, 36).

Phillips, Nathaniel D (2016). *YaRrr! The Pirate's Guide to R*. version: 2017-10-29. URL: www.thepiratesguidetor.com (cit. on p. 60).

Phillips, Webb and Lera Boroditsky (2003). "Can quirks of grammar affect the way you think? Grammatical gender and object concepts." In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 25. 25 (cit. on p. 88).

Picallo, M Carme (2005). "Some notes on grammatical gender and *l*-pronouns." In: *Proceedings of the workshop "Specificity and the evolution/emergence of nominal determination systems in Romance"*. Ed. by Klaus von Heusinger, Georg A. Kaiser, and Elizabeth Stark. Vol. 119. Arbeidspapier, pp. 107–121 (cit. on pp. 71, 96, 108, 128).

Picallo, M Carme (2008). "Gender and number in Romance." In: *Lingue e linguaggio* 7.1, pp. 47–66 (cit. on pp. 71, 96, 108, 128).

Picallo, M Carme (2017). "A note on the locus and function of formal gender." In: *Borealis* 6.1, pp. 1–20 (cit. on p. 72).

Pieters, Tatiana (2020). "A theory of nominal concord." PhD thesis. Brussels, Belgium: Vrije Universiteit Brussel & Université libre de Bruxelles (cit. on p. 63).

Pollock, Jean-Yves (1989). "Verb movement, universal grammar, and the structure of IP." In: *Linguistic inquiry* 20.3, pp. 365–424 (cit. on pp. 181, 193).

Poplack, Shana (1980). "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching." In: *Linguistics* 18, pp. 581–618 (cit. on pp. 25, 26, 30, 31, 42, 51, 171).

Poplack, Shana (2018). *Borrowing: Loanwords in the Speech Community and in the Grammar*. Oxford University Press (cit. on pp. 41, 42).

Poplack, Shana, Alicia Pousada, and David Sankoff (1982). "Competing influences in gender assignment: Variable process, stable outcome." In: *Lingua* 57.1, pp. 1–28 (cit. on pp. 91, 94).

Poplack, Shana, David Sankoff, and Christopher Miller (1988). "The social correlates and linguistic processes of lexical borrowing and assimilation." In: *Linguistics* 26.1, pp. 47–104 (cit. on p. 41).

Prince, Alan and Paul Smolensky (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell Publishing (cit. on p. 43).

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/ (cit. on pp. 60, 115, 159).

RStudio Team (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC. Boston, MA. URL: http://www.rstudio.com/ (cit. on pp. 189, 195).

Rambow, Owen and Beatrice Santorini (1995). "Incremental Phrase Structure Generation and a Universal Theory of V2." In: *Proceedings of NELS*. Vol. 25, pp. 373–387 (cit. on pp. 152, 154, 159, 170, 171).

Riksem, Brita Ramsevik (2017). "Language mixing and diachronic change: American Norwegian noun phrases then and now." In: *Languages* 2.2, p. 3 (cit. on p. 98).

Riksem, Brita Ramsevik (2018). "Language mixing in American Norwegian noun phrases." In: *Journal of Language Contact* 11.3, pp. 481–524 (cit. on p. 98).

Riksem, Brita Ramsevik, Maren Berg Grimstad, Terje Lohndal, and Tor Anders Åfarli (2019). "Language mixing within verbs and nouns in American Norwegian." In: *The Journal of Comparative Germanic Linguistics* 22.2, pp. 189–209 (cit. on pp. 39, 98).

Ritter, Elizabeth (1993). "Where's Gender." In: *Linguistic Inquiry* 24.4, pp. 795–803 (cit. on pp. 69, 72).

Rizzi, Luigi (1990). "Speculations on Verb Second." In: *Grammar in progress: GLOW Essays in honour of Henk van Riemsdijk*. Ed. by Joan Mascaró and Marina Nespor. Groningen: Foris, pp. 375–386 (cit. on p. 145).

Rizzi, Luigi (1997). "The Fine Structure of the Left Periphery." In: *Elements of Grammar*. Ed. by Liliane Haegeman. Springer, pp. 281–337 (cit. on p. 149).

Rizzi, Luigi and Guglielmo Cinque (2016). "Functional categories and syntactic theory." In: *Annual Review of Linguistics* 2, pp. 139–163 (cit. on p. 16).

Roché, Michel (1992). "Le masculin est-il plus productif que le féminin?" In: *Langue française* 96, pp. 113–124 (cit. on pp. 105, 106).

Roodenburg, Jasper and Aafke Hulk (2009). ""Gender" puzzels." In: *Taal en Tongval* 61.1, pp. 143–164 (cit. on pp. 108, 128).

Rooryck, Johan (2003). "The morphosyntactic structure of articles and pronouns in Dutch." In: *Germania et alia. A linguistic webschrift for Hans den Besten*, pp. 1–13 (cit. on pp. 106, 107).

Roy, Jennifer (2019). *I Marie Kondo-Ed a Cluttered Classroom, and Here's What Happened*. Retrieved 20-09-2019. URL: www.weareteachers.com/marie-kondo-classroom/ (cit. on p. 19).

Sande Pineiro, Ariane (2018). "C Plus T As A Necessary Condition For Pro-Drop: Evidence from Code-Switching." PhD thesis. Univeristy of Illinois, Chicago (cit. on pp. 165, 175, 176).

Sanoudaki, Eirini and Guillaume Thierry (2014). "Juggling Two Grammars." In: *Advances in the Study of Bilingualism*. Ed. by Enlli Môn Thomas and Inneke Mennen. Bristol: Multilingual Matters LTD, pp. 216–232 (cit. on p. 46).

Schane, Sanford A (1970). "Phonological and morphological markedness." In: *Progress in linguistics*. Ed. by Karl Erich Heidolph. Berlin: Mouton, pp. 286–94 (cit. on pp. 105, 106).

Schatz Henriette, F (1989). "Code-switching or borrowing?: English elements in the Dutch of Dutch-American immigrants." In: *ITL International Journal of Applied Linguistics* 83-84, pp. 125–162 (cit. on p. 3).

Schütze, Carson T and Jon Sprouse (2014). "Judgment data." In: *Research methods in linguistics*. Cambridge University Press Cambridge, pp. 27–50 (cit. on p. 49).

Schwink, Frederick W (2006). "The reconstruction of variability in Proto-Germanic gender." In: *Variation and reconstruction*. Ed. by Thomas D Cravens. Vol. 4. Amsterdam studies in the theory and history of linguistic sciences. Amsterdam: John Benjamins, pp. 201–212 (cit. on p. 79).

Shim, Ji Young (2013). "Deriving word order in code-switching: Feature inheritance and light verbs." PhD thesis. City University of New York (cit. on p. 38).

Shim, Ji Young (2021). *OV and VO variation in code-switching*. Vol. 1. Current Issues in Bilingualism. Berlin, Germany: Language Science Press (cit. on p. 175).

Sprouse, Jon (in press). "A user's view of the validity of acceptability judgments as evidence for syntactic theories." In: *Linguistic Intuitions*. Oxford University Press (cit. on pp. 49, 50).

Sprouse, Jon and Diogo Almeida (2012). "Assessing the reliability of textbook data in syntax: Adger's *Core Syntax*." In: *Journal of Linguistics* 48.3, pp. 609–652 (cit. on p. 50).

Sprouse, Jon and Diogo Almeida (2017). "Design sensitivity and statistical power in acceptability judgment experiments." In: *Glossa* 2 (1). lingbuzz/001520, pp. 1–32. DOI: https://doi.org/10.5334/gjgl.236 (cit. on pp. 51, 59).

Sprouse, Jon, Carson T Schütze, and Diogo Almeida (2013). "A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001–2010." In: *Lingua* 134, pp. 219–248 (cit. on p. 50).

Stadthagen-González, Hans, Luis López, Maria del Carmen Parafita Couto, and C Alejandro Párraga (2018). "Using two-alternative forced choice tasks and Thurstone's law of comparative judgments for code-switching research." In: *Linguistic Approaches to Bilingualism* 8.1, pp. 67–97 (cit. on pp. 38, 51, 183–185, 194, 198, 199, 227, 229).

Starke, Michal (2009). "Nanosyntax: A short primer to a new approach to language." In: *Nordlyd* 36.1, pp. 1–6 (cit. on p. 18).

Stevens, Stanley S (1946). "On the theory of measurement scales." In: *Science* 103.2684, pp. 677–680 (cit. on p. 54).

Stevens, Stanley S (1951). *Mathematics, measurement, and psychophysics*. Wiley (cit. on pp. 55, 199).

Stowell, Timothy Angus (1981). "Origins of phrase structure." PhD thesis. Massachusetts Institute of Technology (cit. on pp. 180, 181, 194).

Taeldeman, Johan (1980). "Inflectional aspects of adjectives in the dialects of Dutch-speaking Belgium." In: *Dutch Studies*. Springer, pp. 223–245 (cit. on pp. 81, 82).

Thomas, Craig, Carter Bays, and Pamela Fryman (2012). "The final page: Part one (Television series episode)." In: *How I Met Your Mother*. Columbia Broadcasting System (CBS) (cit. on p. 19).

Timm, L A (1975). "Spanish-English Code Switching: El Porqué y How-Not-To." In: *Romance Philology* 28.4, pp. 473–482 (cit. on pp. 30, 31, 51, 156).

Treffers-Daller, Jeanine (1993). "Mixing Two Languages: French-Dutch Contact in a Comparative Perspective." In: vol. 9. Topics in Sociolinguistics. Berlin: Mouton de Gruyter (cit. on p. 26).

Treffers-Daller, Jeanine (2015). *Do balanced bilinguals exist? A critical review of language dominance*. Keynote address at ISTAL 22 (22–24 April). Aristotle University, Thessaloniki (cit. on p. 6).

Trudgill, Peter (2000). *Sociolinguistics: An introduction to language and society*. Penguin UK (cit. on p. 29).

Tsimpli, Ianthi Maria and Aafke Hulk (2013). "Grammatical gender and the notion of default: Insights from language acquisition." In: *Lingua* 137, pp. 128–144 (cit. on pp. 127, 128).

Tucker, G Richard (1999). "A Global Perspective on Bilingualism and Bilingual Education." In: *ERIC Digest* (cit. on p. 1).

Ulutas, Süleyman (2009). "Feature inheritance and subject Case in Turkish." In: *Essays on Turkish Linguistics: Proceedings of the 14th International Conference on Turkish Linguistics, August 6-8, 2008*. Vol. 79. Otto Harrassowitz Verlag, p. 141 (cit. on p. 175).

Vachek, Josef (1964). "Notes on gender in Modern English." In: *A Linguistica*. Vol. 12. Sborník Prací Filosofické Fakulty Brněnské University, pp. 189–194 (cit. on p. 80).

Valdés Kroff, Jorge R, Rosa E Guzzardo Tamargo, and Paola E Dussias (2018). "Experimental contributions of eye-tracking to the understanding of comprehension processes while hearing and reading code-switches." In: *Linguistic Approaches to Bilingualism* 8.1, pp. 98–133 (cit. on p. 46).

Valenzuela, Elena et al. (2012). "Gender and heritage Spanish bilingual grammars: A study of code-mixed determiner phrases and copula constructions." In: *Hispania*, pp. 481–494 (cit. on p. 93).

Van Dulm, Ondene (2007). "The grammar of English-Afrikaans code switching: A feature checking account." PhD thesis. Radboud Universiteit Nijmegen (cit. on pp. 3, 51, 150, 151, 153, 166, 170, 173, 174, 176, 184, 188, 198).

Van Dyck, Tom, Michiel Devlieger, and Michel Vanhove (2009). "Aflevering 3 (Television series episode)." In: *Van Vlees en bloed*. Woestijnvis (produced by Leen Lombaert) (cit. on p. 84).

Van Hell, Janet G, Carla B Fernandez, Gerrit Jan Kootstra, Kaitlyn A Litcofsky, and Caitlin Y Ting (2018). "Electrophysiological and experimental-behavioral approaches to the study of intra-sentential code-switching." In: *Linguistic Approaches to Bilingualism* 8.1, pp. 134–161 (cit. on p. 46).

Van Hell, Janet G, Kaitlyn A Litcofsky, and Caitlin Y Ting (2015). "Intra-sentential code-switching: Cognitive and neural approaches." In: *The*

Cambridge handbook of bilingual processing. Ed. by J W Schwieter. Cambridge University Press, pp. 459–482 (cit. on p. 46).

Vanden Wyngaerd, Emma (2012). "Agreement and syncretism in the articles, adjectives and possessive pronouns in Flemish Dutch." class paper for 'Syntax: Current Issues' (ResMA Linguistics) (cit. on p. 82).

Vanden Wyngaerd, Emma (2017). "The adjective in Dutch-French codeswitching: Word order and agreement." In: *The International Journal of Bilingualism* 21.4, pp. 454–473. DOI: 10.1177/1367006916632302 (cit. on pp. 34, 38).

Vaughan-Evans, Awel, Simon P Liversedge, Gemma Fitzsimmons, and Manon W Jones (2020). "Syntactic co-activation in natural reading." In: *Visual Cognition* 28.10, pp. 541–556 (cit. on p. 46).

Vogt, Hans (1954). "Review of *Languages in Contact*." In: *Word* 10, pp. 79–82 (cit. on p. 29).

Weinreich, Uriel (1953). *Languages in contact: Findings and problems*. Publications of the Linguistic Circe of New York 1. Walter de Gruyter (cit. on pp. 29, 30, 41).

Westergaard, Marit R and Øystein A Vangsnes (2005). "*Wh*-questions, V2, and the Left Periphery of Three Norwegian Dialect types." In: *Journal of Comparative Germanic Linguistics* 8, pp. 117–158 (cit. on p. 149).

Wickham, Hadley (2016). *ggplot2: Elegant Graphics for Data Analysis*. R package. Springer-Verlag New York. ISBN: 978-3-319-24277-4. URL: https://ggplot2.tidyverse.org (cit. on p. 115).

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller (2020). *dplyr: A Grammar of Data Manipulation*. R package version 1.0.0. URL: https://CRAN.R-project.org/package=dplyr (cit. on pp. 115, 159, 189).

Wickham, Hadley and Lionel Henry (2020). *tidyr: Tidy Messy Data*. R package version 1.1.0. URL: https://CRAN.R-project.org/package=tidyr (cit. on pp. 115, 159, 189).

Wilkinson, Leland and Task Force on Statistical Inference (1999). "Statistical methods in psychology journals: Guidelines and explanations." In: *American psychologist* 54.8, p. 594 (cit. on p. 59).

Williams, Cen (1996). "Secondary education: teaching in the bilingual situation." In: *The language policy: Taking stock*. Vol. 12. 2. Llangefni, pp. 39–78 (cit. on p. 27).

Woolford, Ellen (1983). "Bilingual Code-Switching and Syntactic Theory." In: *Linguistic Inquiry* 14.3, pp. 520–536 (cit. on pp. 1, 34, 35).

Zwart, Jan Wouter (1993). “Dutch Syntax: a Minimalist approach.” PhD thesis. Groningen: Rijksuniversiteit Groningen (cit. on p. 149).

Zwart, Jan-Wouter (1997). *Morphosyntax of Verb Movement: A Minimalist Approach to to the Syntax of Dutch*. Vol. 39. Springer Science & Business Media (cit. on p. 145).

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede and Ivo Pletikosić. The style was inspired by Robert Bringhurst's seminal book on typography *The Elements of Typographic Style*.
Cover image: Freepik.com

Final Version as of January 18, 2021