UNIVERSITY OF THE BASQUE COUNTRY (UPV/EHU)

DOCTORAL DISSERTATION

MINIMIZING DEPENDENCIES ACROSS LANGUAGES AND SPEAKERS.
Evidence from Basque, Polish and Spanish and native and non-native bilinguals

---

Author:
Idoia ROS

Supervisor:
Dr. Itziar LAKA

A dissertation submitted in fulfilment of the requirements for the degree of Doctor of Linguistics

at the

Department of Linguistics and Basque Studies

Abstract

Within the last years, evidence for a general preference towards grammars reducing the linear distance between elements in a dependency has been accumulating (e. g., Futrell, Mahowald, and Gibson, 2015b; Gildea and Temperley, 2010). This cognitive bias towards *dependency length minimization* has been argued to result from communicative and cognitive pressures at play during language production. Although corpus evidence supporting this claim is quite broad insofar as grammaticalized structures are concerned (Futrell et al., 2015b; Liu, 2008; Temperley, 2007, among others), its validity rests on more shaky foundations regarding production preferences (Stallings, MacDonald, and O'Seaghdha, 1998; Wasow, 1997b; Yamashita and Chang, 2001, among others). This dissertation intends to address this gap. It examines whether *dependency length minimization* is an active mechanism shaping language production preferences, and explores the specific nature of this principle and its interplay with linguistic specifications and architectural properties of the human memory system. In a series of 5 cued-recall production experiments and 2 complex memory span tasks, I investigate the effect of dependency length in modulating production preferences across languages with differing grammatical properties (e.g., head-position and case marking) and across speakers (e. g., natives and non-natives and with variable working memory capacity).

I begin by showing that the preference for short dependencies is better accounted by a general cognitive preference for minimizing the distance across dependents than by conceptual availability. I then show how languages as diverse as Basque, Spanish and Polish tend to choose the communicatively more efficient structures, when there is more than one available alternative to express the same meaning. Crucially, I confirm that there is consistent variation regarding this tendency both across languages and across speakers. I argue that language-specific (e. g., pluripersonal agreement) and general cognitive mechanisms (e. g., word order based-expectations) interact with the preference towards *dependency length minimization*. Also, I show that the degree of communicative efficiency achieved by highly proficient and early non-native bilingual speakers is lower than that reached by their native peers. Finally, I find that the bias towards shifted orders that yield shorter dependencies correlates positively with working memory. Based on these findings, I conclude that there is strong evidence supporting the claim that *dependency length minimization* is a pervasive force in human language production, resulting from a general cognitive constraint towards efficient communication, and also that its strength varies depending on grammatical and individual specifications compatible with information-theoretic considerations.

Acknowledgements

First and foremost, I owe a great debt of gratitude to my supervisor, Prof. Itziar Laka. Itziar has always pushed me to pursue my personal scientific interests, while granting me total intellectual and financial freedom to make my own mistakes. I also want to express my gratitude to all present and past members of my research group: Irene de la Cruz-Pavía, Nerea Egusquiza, Kepa Erdozia... I am especially thankful to Mikel Santesteban, for his input during the work reported in Chapter 2 of this dissertation, and to Adam Zawiszweski, for his help in Experiment 2 and for making my stay at the Adam Mickiewicz University possible. The research group administrative and technical staff, Yolanda Acedo and Edurne Petrirena, deserve special mention. I want to explicitly thank Edurne for keeping the lab from collapsing and doing so with utmost care and kindness. Special thanks need to go to Luis Pastor for indulging my sweet tooth and for his impeccable help with anything bureaucratic. And thanks also to Sergio López-Sancio, for all the days shared, and for his feedback on some parts of this work. The last couple of years would have been less tolerable without your friendship.

I want to express my gratitude the to the Department of Brain and Cognitive Sciences at the University of Rochester, and to Florian T. Jaeger, who welcomed me there during three great months and made me feel at home, while offering me the possibility of being a part of a thriving community of researchers. I am also thankful to Alfons Gregori i Gomis, at the Institute of Romance Studies at the Adam Mickiewicz University in Poznań, for his hospitality during my stay. Last, I wanted to mention the profound impact that, in the wake of the reproducibility crisis in science, a large number of individuals I have never or only very briefly met have had on my current view of the challenges and possibilities that lay ahead for research: Organizations like the Center for Open Science, the Coursera platform, blogs like Retraction Watch, researchers as Shravan Vasishth, Andrew Gelman, Daniël Lakkens and many others. Also, I could not underscore enough the generosity of the various research communities contributing to open source resources like R or LaTeX, who share their knowledge and provide disinterested help all over Internet and social media. This dissertation would have never been possible without their work.

I thank all the institutions that have provided me with financial support. A PhD and a short-term predoctoral scholarhsip from the Mineco, as well as a predoctoral grant from the Euskara eta Eleaniztasunezko Errektoreordea, have all been crucial to fund this dissertation. Also, I am grateful for the external funding (in the form of travel grants and tuition waivers) received from institutions like NYU Abu Dhabi, the European Society for Cognitive Psychology (ESCoP) or the University of Plymouth, among others.

Most of all, thanks to Matthias, who saved me from starvation and who supports me beyond imagination. Thanks to Amaia, Raúl, Rakel and Vicen, for the beers and the laughs. You all might not have given a damn about this dissertation, but that was exactly the point.

To Matthias and Eneko.

To my family: My mum, my sister, my grandmother, and my aunt Beni. In a world not meant for women, their lives are and have been an example of resilience.

For Lucre, may I have found something you might not have known you lost.

Contents

List of Figures

List of Tables

List of Abbreviations

Abs	Absolutive
Acc	Accusative
AoA	Age of Acquisition
AUX	Auxiliary
BDT	Branching Direction Tehory
Erg	Ergative
Dat	Dative
Det	Determinant
DG	Dependency Grammar
DLM	Dependency Length Minimization
Fut	Future
HNPS	Heavy NP shift
IC	Immediate Constituents
ID	Individual Differences
IO	Indirect Object
L1Ba-L2Sp	Basque-L2Spanish
L1Sp-L2Ba	L1Spanish-L2Basque
Loc	Locative
MaOP	Maximize On-line Processing
MiD	Minimize Domains
Neg	Negation
NP	Noun Phrase
PP	Prepositional Phrase
O	Object
OV	Object-Verb (head-final)
PGCH	Performance-Grammar Correspondance Hypothesis
Pl	Plural
RC	Relative Clause
S	Subject
Sg	Singular
Top	Topic
V	Verb
VO	Verb-Object (head-initial)
WM	Working Memory
WMC	Working Memory Capacity

Chapter 1

General Introduction

Languages in the world share an astonishing number of properties, known as absolute and implicational universals (Dryer, 1992, 2011; Greenberg, 1963; Tomlin, 1986, inter alia, but see Evans and Levinson, 2009). The sheer number of attested languages that comply with these properties demands a theoretical explanation, and casts doubts on the idea that universals could be due to random variation, genetic lineage or areal contact (Dunn, Greenhill, Levinson, and Gray, 2011, but see Dryer, 2011). One promising research avenue in the intersection between psycholinguistics and typology is to investigate how these commonalities emerge and extend across languages and whether they respond to general principles in the brain, shared with other cognitive mechanisms (Christiansen and Chater, 2008; Hawkins, 1994). In fact, a number of word order universals are being investigated from this perspective, and have been proposed to be the result of processing constrains and learnability considerations (Christiansen, 1974; Culbertson, Smolensky, and Legendre, 2012; Gibson et al., 2013; Goldin-Meadow, So, Özyürek, and Mylander, 2008; Hawkins, 1994, among others). Through this dissertation, I focus in depth on one of these postulated general processing constraints affecting word order, the tendency to shorten the distance across elements that belong together, semantically and syntactically.

The principle for shortening dependency length has been claimed to shape word order across usage preferences, grammars and language change. Suggestive evidence for this bias comes from quantitative typology, processing and production on-line preferences and miniature artifical languages (see Section 1). Although the exact mechanism that allows this preference to permeate through grammars of the world is still unclear (but see Fedzechkina, Chu, and Jaeger, 2018), the main linking hypothesis connecting on-line preferences and orders licensed by grammars is put forth explicitly in Hawkins' Performance-Grammar Correspondence Hypothesis (henceforth, PGCH) (Hawkins, 1994, 2004a). He argues that word orders that are optimally efficient in processing in a given language are preferred in performance and that basic word orders derive from conventionalizations of those performance preferences. This idea is implicitly or explicitly present in most of the research on this topic and can be broken down in three main claims: (i) minimizing dependency length is preferred in processing and production (Gibson, 1998, 2000; Hawkins, 1994); (ii) minimizing

dependency length is preferred across grammars (Ferrer i Cancho, 2004; Futrell et al., 2015b; Liu, 2008, among others); (iii) minimizing dependency length impacts the historical evolution of languages' word orders (Aldai, 2011; Tily, 2010). In this dissertation, I test key claim (i), which states that the preference for short dependencies is a general mechanism active in on-line language production, and thus the result of a general processing preference and not of grammaticalized structural preferences specific to any given (set of) language(s). I do so via 5 cued-recall production experiments where I investigate the effect of dependency length across languages (see Chapters 2 and 3) and speakers (see Chapters 4 and 5). Whether and how the preference for reducing dependency length permeates grammars is beyond the scope of this work.

In Section 1, I first describe the phenomenon of *dependency length minimization* (hereafer, DLM) and I review current evidence supporting the claim that this cognitive bias shapes production preferences, grammars and language change. I first focus on previous experimental production studies on the modulation of syntactic choice by DLM (see Subsection 1.1). Next, I cover large-scale corpus and other type of studies, which bear mainly on the effect of DLM on grammars and on the historical evolution of particular languages (see Subsections 1.2 and 1.3). Last, in Section 2, I turn to the idea that language approaches the ideal characteristics of an optimal code for communication. In this context, I will tackle the claims that DLM is shaped by either memory properties (see Subsection 2.1) or by efficiency considerations (see Subsection 2.2). Last, in Section 3, I describe the aims of this dissertation, and the main research questions addressed in each chapter.

1 A general cognitive bias underlying word order: DLM

The preference for short dependencies in language production has been called *distance dependency minimization* (Liu, Xu, and Liang, 2017), *dependency length minimization* (Fedzechkina et al., 2018; Futrell et al., 2015b; Temperley, 2008; Temperley and Gidea, 2010, among others), *Euclidean distance minimization* (Ferrer i Cancho, 2004) or *domain minimization* (Hawkins, 2004a)[1]. In this dissertation, I will use the term *dependency length minimization* (hereafter, DLM), on the grounds of its being more widespread in the relevant literature. Curiously, this situation in the field of language production contrasts with the terminological situation in comprehension, where increasing processing difficulty as a function of the distance between related linguistic elements is referred to simply as (dependency) *locality* effects (Gibson, 2000). In any case, despite the multiplicity of names (that indicates that DLM is an umbrella term for a group of related psycholinguistic accounts), all the aforementioned

[1]For a review on the diffferences between the terms *distance* and *length*, see Liu et al., 2017, where he argues that the term *distance* better reflects language processing as a dynamic and psychological process. Ferrer i Cancho, 2004 uses the term *Euclidean distance*, and focuses on the physical distance between syntcatically linked words. Unlike in DG conventions, both the posistion of the words and their distance to the relevant head - or *sender*, in Ferrer i Cancho's terminology - are of interest.

researchers share the crucial idea that speakers and languages favor short spans across elements bound in a dependency, that is in a syntactic relation.

In natural languages, hierarchical structures or dependencies need to be compressed to fit through a linear and temporally-bounded signal, comprised of sounds or signs in oral and signed languages. One common assumption in psycholinguistics is that processing difficulty arises when these two dimensions are not well attuned, as in the case of long dependencies. Dependencies are linguistic units (frequently, just words) connected to each other by directed links (Tesnière, 1959). Dependency distance or length has been originally defined as "the distance between words and their parents (heads), measured in terms of intervening words" (Hudson, 1995: 16; see also Heringer, Strecker, and Wimmer, 1980; Kimball, 1973: 187 and Temperley, 2008). In research in language production, complexity metrics are usually adopted from a dependency approach to syntax (but see also Hawkins, 1994, where phrase structure grammar is used). Throughout this dissertation, I assume the formalism adopted by dependency grammar (henceforth, DG), due to their relevant role in psycholinguistics, quantitative typology and computational linguistics, where it is used in developing parsing systems for natural languages, as well as in generating treebanks. Among the different conventions used to represent dependencies in DG, arrow arcs are usually favored by Word Grammar (Hudson, 1991). I will adopt this representational convention due its widespread use in the aforementioned fields.

Robust findings in comprehension demonstrate that increasing distance among dependents increases processing difficulty both in subject- and object-extracted relative sentences (Gibson, 1998, 2000) and in other type of sentences (Grodner and Gibson, 2005), as signaled by reading times and other behavioral measures such as re-fixations and regressive saccades in eye-tracking experiments (Bartek, Lewis, Vasishth, and Smith, 2011; Grodner and Gibson, 2005, among others). For example, sentence 1a below entails a smaller cognitive cost than its counterpart 1b (taken from Grodner and Gibson, 2005). This cost is attributed to memory decay and/or interefence at the point of memory retrieval, that is at the most local head (*sent* and *supervised*), as a function of its distance from the target of the retrieval, the dependent (*reporter* and *nurse*) (Gibson, 1998, 2000; Gordon, Hendrick, and Johnson, 2004; Lewis, Vasishth, and Van Dyke, 2006; McElree, 2000; Van Dyke and McElree, 2011). That is, processing difficulty is highly determined by the interplay between the temporal demands of the linearization of hierarchical structures and the properties of the human memory system. The cognitive facilitation derived from reducing the distance across dependents has been claimed to be present also in language production, since orders with shorter spans across dependents are more frequently attested across and within languages (see Subsections 1.1 and 1.2 below for relevant evidence).

(1) a. The *reporter* who *sent* the photographer to the editor hoped for a story
    b. The *reporter* who the photographer *sent* to the editor hoped for a story

It is worth noting that, in language production, researchers initially focused on the impact of categorical definitions of length, also called *weight* or *heaviness* (as in Wasow, 2002, and in previous descriptions in German and English like Behaghel, 1909/10 or Quirk, Greenbaum, Leech, and Svartvik, 1972). This means that the main variable of interest in these works was the relative length or syntactic complexity of the multiple constituents involved in shifting in truth-conditional structures where syntactic choice is available, such as in Heavy NP shift (Ross, 1967), extrapositions (Wasow, 2002), the verb-particle construction (Lohse, Hawkins, and Wasow, 2004) and the dative and locative alternations (Bresnan, Cueni, Nikitina, and Baayen, 2007; Wasow, 2002). The initial pervasiveness of this perspective when tackling length-triggered word order preferences was probably due to the consensus model adopted in the field of language production at the time (for an overview, see Konopka and Brown-Schmidt, 2014), where an extremely incremental scope of planning, highly synchronized with articulation, was assumed to be the main determinant of grammatical encoding.

Throughout this dissertation, I focus on dependency length as the relevant variable of interest, instead of on constituent length or weight[2]. Also, in Chapter 2, I pit against each other the theoretical views and predictions from availability-based models and Hawkins' *PGCH*, for which the distinction between dependency and constituent length is relevant. Dependency length allows a more parsimonious view of converging production and comprehension phenomena, and provides us with a principle of language processing with great predictive power. Also, shifting the perspective from constituent towards dependency length opens up new research avenues on the relation between performance, grammar and historical language change. Moreover, there is growing experimental evidence suggesting that thinking in terms of dependency length is a more adequate way of investigating the preference for certain word orders[3]. In fact, several studies indicate that word order rearrangements are better predicted by relative and not absolute length, that is, by the distance between co-dependents instead of by the specific characteristics of any of constituents involved in shifting alone (Rickford, Wasow, Mendoza-Denton, and Espinoza, 1995: 111; Hawkins, 1994; Stallings and MacDonald, 2011). In Hawkins' English and Japanese corpus data (1994), each additional word added to the longer PP resulted in a higher proportion of shifting, culminating in a 100% *short-before-long* preference for large word differentials between constituents. More relevantly, Stallings and MacDonald (2011) have corroborated experimentally the importance of relative length. When two conditions were matched for relative length but varied in absolute length, there was no effect on ordering

[2]Other measures apart from word or syllable count have been employed in previous literature, such as syntactic complexity (phrasal nodes) (F. Ferreira, 1991; Hawkins, 1994), number of referents (discourse complexity) (Gibson, 1998, 2000), phonological complexity (number of primary stressed syllables) (Anttila, Adams, and Speriosu, 2010; Zec and Inkelas, 1990) or syllabic weight (McDonald, Bock, and Kelly, 1993). Most of these measures have been shown to be highly correlated (Szmrecsányi, 1999; Wasow, 1997b, 2002). In most studies, as in this dissertation, word count is used as an adequate proxy for any of them, due to its easier implementation.

[3]I fully acknowledge that word order variation needs to be explained using a multifactorial approach, as argued by Wasow (1997a) or MacDonald (1993: 196), and that pragmatic, prosodic and other factors play a key role in it (for a detailed discussion, see Wasow, 2002)

preferences. This is the expected outcome if the underlying cause behind certain word orders is minimizing the distance across phrasal heads as claimed by Hawkins (1994) (see Subsection 1.1 below).

1.1 The effects of DLM on syntactic choice

Experimental and small corpus studies provide evidence for a *short-before-long* and a *long-before-short* word order preference as a function of constituent length in head-initial and head-final languages respectively. Crucially, in languages with opposing head-direction, *short-before-long* and *long-before-short* orders minimize dependency length in comparison to their canonically ordered counterparts (see Figure 1.1 below). The goal of most ot the studies reviewed below was testing the claim that constituents occur in order of increasing weight (*End Weight* principle, see Behaghel, 1909/10, *Das Gesetz der Wachsenden Glieder*, and Uszkoreit, 1987, among others), and thus constituent length was manipulated. Nevertheless, in all cases reported here and thourough this dissertation, altering the complexity or length of one of the constituents always entails longer distances between dependents, as can be seen in Figure 1.1 below and in the set of examples 7a - 8b in Chapter 2. Dependency length was thus experimentally manipulated in all the studies reported below, even when that was not originally their main aim (but see Hawkins, 1994). A more detailed overview of these studies will be provided in Chapter 2).

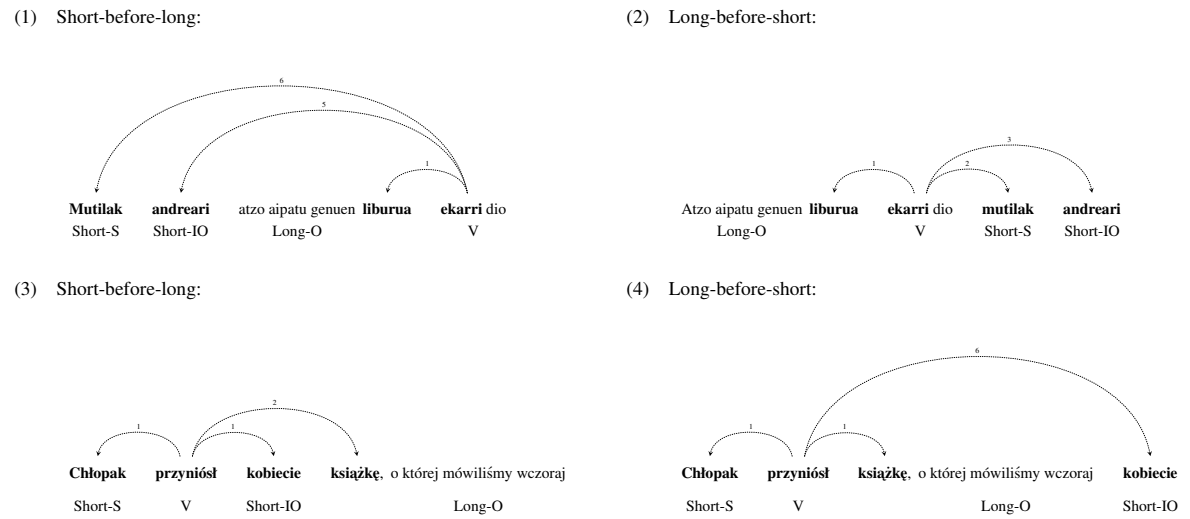


FIGURE 1.1: ***Short-before-long* and *long-before-short* orderings in a head-final (Basque, 1-2) and a head-initial language (Polish, 3-4)**. Non-canonical *long-before-short* in Basque (1) and *short-before-long* in Polish (3) yield shorter dependency lengths than their canonical counterparts (2 and 4). Dependency lengths are here represented through arrow arcs, based in assumptions from DG. For a different type of metrics based on phrasal structure, see Section 1.3 in Chapter 2. (Sentence translation: *The boy brought the lady the book that we mentioned yesterday*)

The tendency to linearize long and complex constituents after short and simple ones has been observed in German as early as 1909 (Behaghel, 1909/10, *Das Gesetz der Wachsenden Glieder*, and Uszkoreit, 1987). The same preference was oberved also for English (Quirk et al., 1972: 148, Ross, 1967) or Dutch (Haeseryn, 1997). Later, the preference for *short-before-long* orders has been attested in small corpus studies (between 5,016 sentences and 210 pages long, depending on the language) in English (Hawkins, 1994; Wasow, 2002) and, to some extent, in a few other head-initial languages such as Hungarian, Rumanian, German, Polish, Greek and Finnish (Hawkins, 1994). This preference has also been found for English in 2–5 year old children (de Marneffe, Grimm, Arnon, Kirby, and Bresnan, 2012) and in several experimental studies (Arnold, Losongco, Wasow, and Ginstrom, 2000; Stallings et al., 1998 and Stallings and MacDonald, 2011).

Unlike in head-initial languages, where speakers favor a *short-before-long* tendency when shifting sentence constituents that vary with respect to their length, the opposite *long-before-short* preference has been found in head-final languages. This tendency has been observed in small corpus studies carried out by Hakuta (1981), Hawkins (1994), Kondo and Yamashita (2011) and Choi (2007). Choi (2007) studied word order variations in the dative-accusative alternation in Korean using a spoken million word corpus. Hakuta reported the *long-before-short* preference in 3–6 year old Japanese children. Also, Kondo and Yamashita, 2011 report a preference for *long-before-short* orders in transitive and ditransitive sentences in spontaneous speech in Japanese. Hawkins (1994) used small Japanese, Korean and Turkish written corpora (230 to 91 pages of data). In these three head-final languages, he found that speakers place long constituents before short ones. Three production experiments have confirmed the *long-before-short* word order preference in Japanese, Korean and Persian (Dennison, 2008; Faghiri, Samvelian, and Hemforth, 2014; Yamashita and Chang, 2001).

1.2 The effects of DLM on the form of grammars

Unlike in the detailed corpora and experiments mentioned in Section 1.1 above, where syntactic choice in a limited number of constructions and languages was the main focus of attention, the aim of the large corpus studies that I will review here is to obtain an estimate of the extent to which short dependencies shape grammars and contribute to observed general word order universals. The approach used in these works is to measure either the average or the total sum of dependency lengths in several natural languages in dependency treebanks (that is, syntactically annotated corpora), and to compare it with random and optimized reordered baselines (Ferrer i Cancho, 2006; Ferrer i Cancho and Liu, 2014; Futrell et al., 2015b; Gildea and Jaeger, 2015; Liu, 2008; Temperley and Gidea, 2010, among others). This approach contrast with previous attempts to investigate whether certain universals where derived from processing preferences, in which actual word order distributions attested in the languages of the world were compared to those predicted by theories like Dryer's Branching Direction Theory (BDT) or Hawkins' PGCH (Dryer, 1992; Hawkins, 1983, 1994).

The first study to compare actual grammars with randomized simulations was Ferrer i Cancho (2004). He found that, in Romanian and Czech treebanks, original dependency lengths were shorter and increased slower than when trees were randomly scrambled. Later, in Gildea and Temeperley (2010), dependencies from four German and English corpora were extracted and compared with both optimized and random baselines (see also Park and Levy, 2009). Both languages showed an effect of DLM, although stronger in English than in German, which scored closer to its optimized baseline than the latter (see also previous work exclusively in English in Temperley, 2007 and Temperley, 2008). Crucial evidence for an above chance bias towards DLM is provided by two large-scale crosslinguistic studies conducted by Liu (2008) and Futrell et al. (2015b), where a total of 35 typologically diverse languages were included[4]. All these languages present shorter dependency lengths than their randomized baselines. Also, in Futrell et al. (2015b), original orderings were compared with an optimized baseline for the minimum possible dependency length in a projective baseline, to observe how far from the ideal in terms of DLM each language would score (see also Gildea and Temperley, 2010).

Methodological divergences between these four works deserve further mention. Reordered dependencies can either be kept constant across comparisons (the original tree structures are maintained, and only word order is shuffled) as in Futrell et al. (**)** or GIldea and Temperley (2010) or not (both word order and tree structure are reordered, while maintaining the same number of nodes and the same lexical elements) as in Liu (2008) or Ferrer i Cancho (2004). Also, Liu (2008) created two random baselines, one constrained by projectivity, where dependency arcs did not cross, and another one where crossing arcs were allowed. In contrast, in Futrell et al. (2015b), all possible grammars were projective, and sister nodes where either randomly permuted while maintaining head direction given the type of dependencies involved (fixed order baseline) or not (free word order baseline). These differences carry significant implications. As Futrell, Mahowald, and Gibson (2015a) argue, conservative baselines where tree structure, projectivity and consistency in head direction are maintained provide more direct and robust evidence for the effect of DLM on syntax. In contrast, the results from Liu (2008) are consistent with other mechanisms that can help reducing dependency length besides word order, such as altering what content to express or reducing the number of constituents (see Hiranuma, 1999; Pastor and Laka, 2013; Ueno and Polinsky, 2009). Nevertheless, projectivity and head consistency are claimed to be derived from the principle of DLM (Ferrer i Cancho and Díaz-Guilera, 2007; Ferrer i Cancho, 2015). This might be a correct hypothesis, but causal evidence for it is still lacking. In any case, if projectivity, head consistency and DLM should all be independent forces behind word order distributions, the baselines in Liu (2008) would be weaker evidence for the role of DLM in shaping language form than in the case of Futrell et al. (2015b).

[4]Liu (2008) investigated the bias towards DLM using dependency treebanks of the following twenty languages: Arabic, Basque, Bulgarian, Catalan, Chinese, Czech, Danish, Dutch, English, German, Greek, Hungarian, Italian, Japanese, Portuguese, Romanian, Slovenian, Spanish, Swedish, Turkish. Futrell et al. (Futrell2015a) include 14 languages more: Ancient Greek, Bengali, Estonian, Finnish, French, Hebrew, Hindi, Indonesian, Korean, Latin, Persian, Russian, Slovak, Tamil and Telugu

Gildea and Jaeger (2015) controlled the role of lexical predictability, while measuring the preference for DLM in English, German, Czech, Chinese and Arabic dependency treebanks. They measured average and per sentence dependency length and average log-probability of the next word given the two previous ones in randomly created fixed-word order baselines, and compared them with the original grammars. The ratios obtained in all these languages were lower than expected by chance in terms of dependency length and higher in terms of average log probability. They also find that languages tend to be around some optimal frontier that trades off dependency length against log probability.

Last, some researchers have proposed wider effects of the preference for short dependencies on language form. Hawkins (1994) identifies certain areas of grammar for which a preference for adjacency of phrasal heads (and, hence, shorter dependencies) is claimed to be the underlying principle such as basic word order distributions, wh-movement and hierarchies of relativization. Ferrer i Cancho (2006) argues that the exceptionality of dependency crossings in natural languages is a side effecy of minimizing the distance between syntactically related words (see also Ferrer i Cancho and Díaz-Guilera, 2007 and Ferrer i Cancho, 2015). Similarly, Temperley (2008) shows that dependency length is reduced when branching is consistent and opposite branching of certain one-words is allowed, and argues that this preferences impact grammatically licensed orders in English (see also Temperley, 2007). In this sense, both consistency and inconsistency of head direction (Greenberg, 1963; Hawkins, 1994: 98-100 (Universals 2-4), Dryer, 1992; Lehman, 1973, among others) would be modulated by DLM.

1.3 The effects of DLM on language change

A couple of corpus and one artificial miniature language study provide suggestive evidence that DLM could be a driving force behind the historical evolution of the grammars of natural languages. Tily (2010) compared random and optimized reorderings of above 150,000 sentences covering the years 850-1500 in Enlish, and found that Old English yielded higher values for average dependency lengths than more recent English, while controlling for differences between genres, dialects and texts. In other words, dependency lengths grew shorter over time in Old and Middle English when compared with random reordering of the same sentences. Tily links this preference to the word order change from OV to VO undergone by English during that period. Compellingly, his study shows that while dependency lengths decrease between verbs and their arguments, they remain roughly constant over time between elements internal to NPs, where English underwent no diachronic change. In turn, Aldai (2011) addresses the word order change that *wh-phrases* have sustained in Basque from the 16th to the 20th century. In a corpus analysis of the Gospels of Matthew and John, Aldai (2011) found that, while in the 1571 data 59.7% and 40.3% of the *wh-phrases* and the verbs were placed in adjacent and non-adjacent position respectively, this tendency reversed in his 1999 database, with 97.1%, and 2.9% of adjacent and non-adjacent constructions respectively.

Namely, out a of a situation with more flexible word order in the 16th century, the distance of the fronted *wh-phrase* and the verb relative to each other has been reduced and grammaticalized. These results are highly compatible with the hypothesis that grammars evolve either to maintain or to achieve more efficient linearizations, namely those with shorter dependency lengths.

Lastly, Fedzechkina et al. (2018) found that, when exposed to a miniature language with long dependencies, learners restructured this input and reduced dependency lengths, even when this required employing a *long-before-short* linearization opposed to the one used in their native language. These results are interpreted as compelling evidence for a casual link between individual on-line preferences during language production and the distributional patterns observed across languages in the world, which tend to reduce dependency length. These deviation in the newly acquired input, caused by general cognitive bias, could act as a seed in language change if they accumulate and spread over generations (for similar proposals, see Kirby, Cornish, and Smith, 2008).

In sum, evidence for DLM as a principle shaping language production preferences and grammars stems mainly from results from three sources, namely behavioral experiments, text frequency counts in small hand-coded corpus studies and quantitative typological data from large-scale treebanks. Results from different methodologies all show a general bias for short dependencies and for mirror-like *short-before-long* and *long-before-short* nesting preferences of constituents of varying length in head-initial and head-final languages. Furthermore, there is initial evidence compatible with the idea that languages evolve to reach efficient processing configurations, and at least one study gives a hint of at some plausible cognitive mechanisms behind this tendency (Fedzechkina et al., 2018). Nevertheless, as I point out in Section 3 below, where I lay out the aims of this dissertation, further research is much needed. Research needs to include more typologically diverse languages, while employing a more fine-grained taxonomy of possible basic word orders (see Aldai, 2011: 1113). In addition, we need to take into consideration the role of language-specific architectural traits such as case marking or agreement in efficient language use, since both devices play an important part in the processing of linguistic structures. Last, research on the underlying causes for the preference for short dependencies in production is vital, since so far this issue has never been directly addressed.

2 Efficient communication, memory and DLM

It has been recently claimed that long dependencies are costly, and thus avoided, not only in comprehension, but also in language production (see Futrell et al., 2015b; Liu, 2008; Scontras, Badecker, Shank, Lim, and Fedorenko, 2015; Temperley and Gidea, 2018, but see MacDonald, Montag, and Gennari, 2016). In contrast, other (or sometimes the same) researchers have linked

the preference for DLM to communicative efficiency (Futrell et al., 2015b; Gildea and Jaeger, 2015; Hawkins, 1994). In the next two sections, I briefly review current perspectives on the underlying causes of the bias towards DLM.

2.1 DLM as the result of memory constraints

In language comprehension, working memory constraints have long been posited as the main underlying cause for the preference to reduce dependency lengths across co-dependents, due to activation decay or encoding and retrieval interference (Gibson, 1998, 2000; Lewis and Vasishth, 2005; McElree, 2000; Nairne, 2002; Van Dyke and Lewis, 2003; Van Dyke and McElree, 2011, among others). Language comprehension unfolds in time and is thus necessarily highly incremental. Upcoming words are integrated one by one to the ongoing representation of the sentence in question. Sometimes retrieval of previous words is required to access a correct intepretation, which might be hindered by increasing temporal and linear distance of the relevant material with respect to the retrieval point. Whatever the exact implementation mechanisms, there is little controversy on the role of memory as the driving force(s) behind the decrease in processing difficulty when comprehending long dependencies (for the role of expectation- and contraint-based processing and its interplay with memory limited resources, see Husain, Vasishth, and Srinivasan, 2014; Levy, Fedorenko, and Gibson, 2013; Safavi, Husain, and Vasishth, 2016; Vasishth and Drenhaus, 2011, *inter alia*). Evidence from planned experiments and corpora support this hypothesis (but see also Gennari, Mirkovic, and MacDonald, 2012; MacDonald, 2013), and specific predictions regarding word-by-word location of processing difficulty are met in experimental studies (Grodner and Gibson, 2005; Levy, 2008).

Production research has traditionally assumed that word order preferences are derived from the need to ease production by reducing memory load (Arnold et al., 2000; Bock and Irwin, 1980; V. S. Ferreira and Yoshita, 2003; V. S. Ferreira and Dell, 2000; MacDonald, 2013; Wasow, 2002, among others). Traditionally, in production, difficulty is indirectly measured by relative frequency when more than one truth-equivalent structure are available: The most frequent structure is assumed to be the one that is easier to produce[5]. In the same vein, word order preferences have been attributed to the need to maintain fluency, as a consequence of delayed lexical and phonological planning due to similarity-based interference when trying to retrieve inteded items (Gennari et al., 2012; Humphreys, Mirković, and Gennari, 2016, *inter alia*). It has been proposed that easy to retrieve elements due to a higher accessibility such as

[5]Word latencies and durations are also assumed to be indexes of production difficulty. Scontras et al. (Scontras et al., 2015) find shorter word durations and latencies for object- versus subject-extracted relative sentences and wh-clauses with longer versus shorter filler-gap dependencies, which they interpret as evidence for their higher cognitive cost due to increased dependency distance between dependents (but see MacDonald et al., 2016). While word latencies can be a reliable signature of the time needed to plan a word and thus can be understood as a sign of production difficulties, it is unclear that word durations can be interpreted this way (see Jaeger and Tily, 2010).

short constituents are produced earlier (Arnold et al., 2000; de Smedt, 1994; V. S. Ferreira, 1996; MacDonald, 2013; Stallings et al., 1998, among others).

More recently, the preference for DLM has been related to a higher cognitive cost of long dependencies (see Section 2.2 below). Shortening dependency lengths has been connected with limitations in working memory since Yngve's (1960: 452) *Depth Hypothesis*, which posits a limit to grammars and use that does not exceed "a number equal or near equal to the span of immediate memory presently assumed a 7±2". Recently, other researchers have proposed that the preference for DLM emerges from the interaction of the incremental nature of the production system and the properties of the human memory system (Futrell et al., 2015b; Jaeger, 2010; Liu, 2008; Temperley and Gidea, 2018, inter alia). Futrell, for example, argue that "because storing items in memory may be difficult or error prone, short dependencies would be easier and more efficient to produce and parse" (Futrell et al., 2015b: 10336). I will review Futrell's proposal below, since memory cost is assumed to be a processing constraint affecting the preference for *information locality* derived from considerations of communicative utility, appealing to both memory and efficiency considerations. Liu links working memory limits and the average dependency lengths presented in a large sample of human languages. According to this author, as Yngve (1960), "average dependency distance within languages should remain below a threshold", which they "expect to be smaller than 4" because that number should signal the limits of working memory capacity (Liu, 2008: 13). However, the sources and loci of production difficulty in the case of DLM are not as explicitly laid out as in language comprehension. First, little is known about the *chunks* or types of representation shared by language production and working memory mechanisms. Also, it is still unclear why working memory should push language towards a system with short dependencies, given that language production does not necessarily work highly incrementally.

2.2 DLM as the result of efficiency considerations

The most puzzling aspect of the bias towards short dependencies is that it is present in both comprehension and production, despite their differences. Short dependencies are advantageous for the listener, and preferentially produced by the speaker. This converging evidence is in line with the increasingly influential claim that language form is shaped by requirements of efficient communication. This hypothesis is clearly at odds with the position that language for communication might be a "kind of epiphenomenon", since "for that purpose [communication], language is not well designed, because you have such properties as ambiguity." (Chomsky, 2002: 107, as cited by Piantadosi, Tily, and Gibson, 2012, where they show that ambiguity is advantageous for in-context communication; see also references therein). In this section, I broadly sketch the claim that language form and linguistic preferences are shaped by communicative efficiency, which requires an equilibrium between informativeness and simplicity, and tackle the issue of how the bias towards DLM has been proposed to emerge from this key property of language.

Since as early as the beginning of the twentieth century, functionalist accounts have attempted to explain implicational universals and other properties present in languages across the world through the trade-off between competing motivations such as utility and simplicity (Gabelentz (1901: 256), as cited by Regier, Kemp, and Kay, 2015; see also E. Bates, McNew, MacWhinney, Devescovi, and Smith, 1982; Comrie, 1989; Croft, 2001; Dryer, 1992; Haspelmath, 1999; Hockett, 1960; Slobin, 2001, among others). Nevertheless, these attempts at uncovering the role of communicative goals in shaping language form lacked falsifiable and operational definitions of notions such as complexity and communicative utility. In contrast, current accounts of efficient communication are usually formalized in terms of information theory (Shannon, 1948) or Bayesian inference (Jaeger and Buz, 2017), and employ objective measures of cognitive cost and facilitation provided by current theories of sentence processing (see Jaeger and Tily, 2010). The key idea in information-theoretic approaches to human languages is that language approximates a code for efficient communication (Shannon, 1948; Zipf, 1949). *Efficiency* means that humans make a rational use of limited resources given task's demands and reward (Anderson, 1990). *Rational* means perfectly adaptive in so far as ideal agents act to maximize the expected utility of their behavior given their cost. Languages are thus seen as the result of a utility function, by which agents maximize reward while reducing cost. In other words, speakers try to successfully and reliably convey the intended meaning with the least cognitive effort. This view has been initially supported by evidence stemming from lexical items' lengths (Ferrer i Cancho and Solé, 2003; Piantadosi, Tily, and Gibson, 2011; Zipf, 1949) and durations (Aylett and Turk, 2004) and semantic category systems such as color, numeral systems and kinship (see Regier et al., 2015's review). Recently, syntactic preferences in human languages (Jaeger, 2010; Kurumada and Jaeger, 2015) and when producing gestures (Gibson et al., 2013) have also been shown to be subservient of efficient communication.

The bias towards DLM has been proposed to result from communicative efficiency considerations (Fedzechkina et al., 2018; Fedzechkina, Jaeger, and Newport, 2012; Ferrer i Cancho, 2004; Futrell et al., 2015b; Hawkins, 1994). Specific proposals, however, differ in the content and shape of the utility function, specifically in how cost is factored in. In Hawkins' proposal, distance across phrasal heads is minimized in order to enhance efficiency by reducing processing complexity (Hawkins, 1994, 2004a, 2014, and Frazier, 1985; Kuno, 1974; Miller and Chomsky, 1963; Yngve, 1960 for previous proposals linking word order preferences and processing mechanisms). Complexity is understood as the amount of structure the parser needs to survey to recognize syntactic structure. Word orders that require fewer computations (fewer words, and where dependency length is thus shorter) in order to recognize syntactic and semantic structure are considered as optimally efficient[6]. For Ferrer i Cancho (2004), "the distance between syntactically related items in sentences is a basic ingredient of the cost of a sentence". For Futrell (2017), the cost is the overall entropy of the language itself. In his

[6]In Hawkins' PGCH, word order preferences do not consistently correlate with memory cost, unlike in the accounts reviewed in Section 2.1. While Hawkins acknowledges that memory load may be reduced when dependency length (or *domain minimization* in PGCH terms) is reduced, other grammatical and performance distributions derived from considerations of efficient processing correlate with increases in memory load (see Hawkins, 2004a: 214-5, 234)

account, the preference for DLM is due to information constraints imposed on efficiency requirements by memory limitations in an extremely incremental system. Specifically, Futrell argues that DLM is a subset of the preference for what he terms *information locality*: elements that are dependent either for meaning or for their distribution need to be close together in the speech signal, in order to avoid decay in the effectiveness of contextual cues for predicting words due to noise, which increases linearly along dependency distance (see also Futrell and Levy, 2017). This hypothesis resembles Gildea and Jaeger's proposal (2015: 1-2), which they termed *locality of information*: "Low dependency length and low locally-conditioned information density both indicate that the information in a sentence is arranged locally, such that related words are near one another. That is, cues that are informative about an upcoming word tend to immediately precede the word. Intuitively, such locality of information facilitates efficient sentence processing by allowing integration of information into a complete sentence interpretation while keeping memory load low.". All these accounts face several problems. Hawkins' complexity metrics are not supported by psycholinguistic evidence, which indicates that language comprehension is highly incremental and does not place elements in a buffer until reaching a complete representation. In the case of Ferrer i Cancho, the preference for short dependencies hinges critically on the assumption that language production is also dependent on the impact of memory limitations. Furthermore, Futrell adds a second unsubstantiated assumption about the nature of incrementality in the production system. As yet, it is premature to transfer to production what we know about the architectural principles at work in comprehension and about its interaction with other cognitive systems. The degree of incrementality in language production is strategically controlled, and differences in the scope of planning are dependent on factors such as language, type of construction, time pressures, etc. (for a review, see **??**). Furthermore, before articulation, linguistic representations are hierarchical, and it is unclear how and at what stage these representations interact with retrieval mechanisms.

In this dissertation, I assume that long dependencies are disfavorable for language comprehension, but, unlike in the accounts above, not necessarily for production. If language is shaped by considerations of efficiency, speakers will try to produce short dependencies, to ensure successful and robust communication of their message by reducing any possible reconstruction error on the part of the listener. I will make no claims on whether this is due to some type of *audience design* (that is, explicitly tailoring utterances to the need of the listeners, see H. H. Clark and Murphy, 1982), the result of implicit learning via explicit or implicit feedback about the success of previous communication intents involving long dependencies (Jaeger, 2013) or whether it happens by means of some other type of learning or cognitive mechanisms (Fedzechkina et al., 2018). My proposal differs from previous ones in that I attribute cost to the production of shifted word orders, which are used to reduce dependency length, and not to the production of long dependencies per se. This second assumption is supported by evidence from Myachykov, Thompson, Scheepers and Garrod (2011) and Myachykov, Scheepers, Garrod, Thomspn and Fedorova (2013) and further research is needed to confirm that these assumptions are correct. The trade-off between these two competing

motivations makes a series of predictions that are met in the results presented in this dissertation (see Chapters 4 and 5). The expected contribution of this dissertation does not crucially hinge in this set of assumptions.

3 Aims and outline of the dissertation

In this section, I outline the main research questions addressed in the following four chapters (see Chapters 2 to 5), where I sought to gather relevant evidence in order to either support or falsify the claim that DLM is shared across languages (**Part 1. DLM across languages**, comprising of **Chapters 2** and **3**) and individuals (**Part 2. DLM across speakers**, comprising of **Chapters 4** and **5**). Specifically, I investigated word order preferences as a function of dependency length by means of 5 cued-recall production experiments in Basque, Polish and Spanish, where native and non-native speakers and individuals with varying working memory capacity participated. Cross-linguistic research is of vital importance in psycholinguistics, whose main aim is to build an overarching theory of human language production and processing that applies to any possible language. Also, research on individual differences has been mainly absent from psycholinguistics and linguistic research (see Kidd, Donnelly, and Christiansen, 2018) and it is also of relevance for this goal. Nevertheless, current psycholinguistic theories in language production are still based on evidence from less than 1% of the world languages (see Jaeger and Norcliffe, 2009 and Norcliffe, Harris, and Jaeger, 2010, for relevant reviews), most of them belonging to the Romance and Germanic branches of the Indo-European family, and individual differences across speakers are also hardly addressed. Methodologically, I expect this dissertation to contribute towards the goal of correcting these shortcomings.

- In **Chapter 2**, I used Basque to explore the possibility that the preference to minimize dependency distance is not driven by the need to facilitate processing, but to ease production, as claimed by availability- and other related production accounts. Basque is a flexible language that allows pre- and postverbal placing of the dependents in relation with the main verb. So far, only rigidly head-final languages like Korean or Japanese have demonstrated a *long-before-short* preference when producing sentences containing constituents of different length. Although Persian is reported to allow some constituents in postverbal position in colloquial speech, such as clauses and goal-denoting arguments, verb position is not addressed in Faghiri and Samvelian, 2014 nor in Faghiri et al., 2014, and only verb-final orders are reported in both studies. I explored whether Basque presents this same *long-before-short* preference, which yields shorter dependency lengths than either canonical or *short-before-long* orderings. Additionally, I addressed the issue of whether pluripersonal verb agreement, which, unlike Japanese and Korean, Basque presents, contributes to modulating word order preferences when constituents are long and, thus, costly to process. This work was published as Ros, I., Santesteban,

M., Fukumura, K.& Laka, I. (2015). Aiming at shorter dependencies: the role of agreement morphology. *Language, Cognition and Neuroscience, 30(9)*

- Recent corpus research has provided evidence that a wide number of languages has an above-chance preference for short dependency lengths. However, languages seem to vary in the extent of this preference. In **Chapter 3**, I sought to explore whether previously unexplained variance in large and small corpus studies was due to random noise or to differences derived from language-specific properties. To do so, I compared the magnitude of the preference for minimizing dependency length across languages with rich and poorer case morphology (Basque and Polish versus Spanish), and with differing basic word order (Polish and Spanish versus Basque). From an information-theoretical perspective, both word order and case morphology could impact the processing cost of long dependencies and reduce the need to resort to DLM. Furthermore, the only experimental evidence on the preference for *short-before-long* orderings in head-final languages came from English. I replicated this result in Polish and Spanish. This work has been submitted as Ros, I., Zawiszweski, A. & Laka, I. *Cross-linguistic differences in the magnitude of the preference for shortening dependencies as the outcome of efficient communication*.

- In **Chapter 4**, I investigated the bias for DLM in native and non-native speakers of Basque and Spanish, two languages that minimize dependency lengths via mirror-like linearization preferences, *long-before-short* in Basque and *short-before-long* in Spanish. It has been claimed that highly proficient bilingual speakers transfer the linearization favoured in their native language to their non-native one (Dennison, 2008), which casts doubts on the universality of the preference for reducing dependency length and on its being the source of the opposite *long-before-short* and *short-before-long* linearizations. I tested this claim, and gathered more evidence to address the question of whether differences across native and non-native speakers could be better accommodated by the idea that non-native speakers are less apt to adhere to the balanced trade-off between production ease and communicative success inherent to communicative efficiency.

- The preference for short dependencies has been related to either efficiency considerations or to limitations in memory capacity. In **Chapter 5**, for the first time to the best of my knowledge, I addressed the underlying causes under the preference for DLM in language production. To do so, I conducted a dual-task production experiment, along with two tasks developed to measure memory span (Unsworth, Heitz, Schrock, and Engle, 2005), where native speakers of Basque participated. I sought to test the direction of the relationship between memory demands and the preference for short dependencies, by pitting against each other approaches that make opposing predicitons due to their differing views on the preference for DLM as the result of memory limitations or of efficiency requirements.

- In **Chapter 6**, I conclude, summarize the main findings and contributions of this dissertations and sketch possible lines of further research.

Part I

Dependency length minimization across languages

Chapter 2

General and language-specific mechanisms to enhance language efficiency

Abstract

Approaches to quantitative typology and psycholinguistics claim that all languages respond to a processing-based linguistic universal, namely to a general preference for dependency length minimization (DLM), that is for shortening the distance across dependencies. In this chapter I examined word order alternations as a function of dependency length in Basque. Basque is an OV language with flexible sentence word order and rich verb agreement, which can shed light on the universality of this processing principle. Availability models in psycholinguistic research argue that all languages will prefer a *short-before-long* ordering of constituents, independently of their typological configuration and of the resulting increased distance of co-dependents. Contrary to them, efficiency-based accounts predict a preference for short dependencies. In particular, Hawkins (2014) claims that whereas *short-before-long* orders will be favored in VO languages such as English, the mirror-like *long-before-short* orders will be preferred in OV ones such as Japanese. Evidence from rigid OV languages supports this claim. However, it is unclear how dependency length affects word order preferences in an OV language that has pluripersonal verb agreement and allows post-verbal arguments, such as Basque. I found a general *long-before-short* preference and a tendency to place the verb in a sentence-medial position when one constituent is long. I argue that since agreement morphology unequivocally signals the thematic role and case of surrounding phrases, it contributes to speeding up sentence processing. I conclude that morphologically rich

languages employ both general adjacency mechanisms that allow for short dependencies and language-specific resources in order to enhance language efficiency[1].

RESEARCH QUESTIONS:

- Is dependency length minimization a general processing principle, which affects all languages independently of their typological make-up? That is, does Basque, a flexible OV language, present a *long-before-short* preference when ordering constituents relative to dependency length?
- Is there a universal preference for ordering short elements before long ones, as predicted by availability-based accounts? That is, does Basque, a flexible OV language, present a *lshort-before-long* preference when ordering constituents relative to dependency length?
- Aside from language-universal principles, do language-specific mechanisms enhance language processing? That is, does Basque, a language with pluripersonal agreement, employ verb morphology as a way to improve language processing efficiency?

1 Introduction

In the language sciences, the impact of processing demands on sentence word order is a central topic of research. In the last decades, linguistics has come to accept that processing constraints play a role in word order preferences (Chomsky, 1995; Dryer, 1980, 1992; Hawkins, 1983, 1994; Wasow, 2002), thus converging with psycholinguistic models that have tackled the question of how processing demands impact on sentence word order variations (Arnold, Wasow, Asudeh, and Alrenga, 2004; Bock, 1982; Bock and Levelt, 1994; Bresnan et al., 2007; F. Ferreira and Swets, 2002; Gibson, 1998, among others), including accounts based on information theoretic considerations (Gibson et al., 2013; Maurits, Navarro, and Perfors, 2010, *inter alia*). Thus, there is general agreement that sentence word order is modulated by factors external to the grammar, although there are differing views on the specifics of this interaction.

The question of exactly how processing constraints underlie word order preferences in sentence production has remained controversial to date. *Availability-based models* of sentence production claim that word order variations are primarily affected by ease of production, whereas *efficiency-based theories* focus more on ease of communication, that is on a trade-off between production cost and success in conveying the intended message (Hawkins, 2014, *inter alia*). In availability models, speakers favor sentence structures where more readily

[1]The content of this chapter has been partially published as Ros, I., Santesteban, M., Fukumura, K.& Laka, I. (2015). Aiming at shorter dependencies: the role of agreement morphology. *Language, Cognition and Neuroscience, 30(9)*, 1156–1174.

accessible phrases are placed earlier, to reduce the need of holding already-retrieved information in memory and to start producing early (Bock and Irwin, 1980; Bock and Warren, 1985; Branigan, Pickering, and Tanaka, 2008; Bresnan et al., 2007; V. S. Ferreira and Yoshita, 2003; V. S. Ferreira and Dell, 2000; Levelt and Maassen, 1981; Prat-Sala and Branigan, 2000; Tanaka, Branigan, McLean, & Pickering, 2011; see Jaeger and Norcliffe, 2009, for an overview). In contrast, efficiency-based theories argue that word orders that allow the fastest computation of constituent structure are preferred across languages (Hawkins, 1994). Specifically, Hawkins (1994, 2004a, 2014) put forth a model that explains how speakers preferentially choose word orders that can optimally facilitate parsing, which in turn is claimed to influence production preferences, leading to a general preference for linearizations with short dependencies. Other accounts straightforwardly predict a general bias for short dependencies, but not that word order preferences will necessarily vary depending on typological properties such as head direction (whether dependents follow or precede heads in phrases (Futrell et al., 2015b; Gibson, 1998, 2000; Temperley and Gidea, 2010, among others). I will thus not cover these accounts here (but see Chapters 3 and 5).

In this chapter, I explored how phrasal length affects sentence word order in Basque, a heavily inflected non-rigid OV language. I investigated the impact of length (operationalized as ortographic word count) on the production of transitive and ditransitive sentences in Basque, for which availability-based and efficiency-based models make opposing predictions. As I discuss in Section 1.2 in more detail, availability-based accounts predict a general preference to place phrases that are short before those that are long, independently of the basic word order or other typological properties of the language (Arnold et al., 2000; de Smedt, 1994, Easy First principle in MacDonald, 2013; Stallings and MacDonald, 2011; Stallings et al., 1998; Wasow, 1997a, 1997b). In contrast, in Hawkins' *Performance-Grammar Correspondence Hypothesis* (henceforth, PGCH)[2], word order preferences depend on phrasal length and head direction of the language in question. In Hawkins' theory, head-initial (VO) languages will prefer *short-before-long* orders, while head-final (OV) languages will favor *long-before-short* orders[3]. Crucially, these mirror-like word order preferences result from a general preference for dependency length minimization for all languages regardless of typological differences. In availability-based models, in contrast, word orders that allow for ease of production will ultimately lead to longer dependencies in head-final versus head-initial languages (see Section 1.3).

In this chapter, I will explain how investigating Basque can enhance our understanding of the effect of dependency length on shaping language form. In Section 1.1, I first describe the relevant linguistic characteristics of Basque. After that, in Sections 1.2 and 1.3, I contrast the

[2]This account has also received the name of *Performance Theory of Order and Constituency* (Hawkins, 1994). I adopt the most recent denomination (Hawkins, 2004a and Hawkins, 2014).

[3]With the terms *short-before-long* or *long-before-short*, I refer to the relative position of the long and short phrasal constituents with respect to each other, independently of whether they are piled up to one side or placed to both sides of their common head (see Temperley, 2007 for English, Jiang and Liu, 2015 for Chinese and English and Ros, Santesteban, Fukumura, and Laka, 2015 for Basque).

availability-based accounts and the PGCH model, and their respective predictions for Basque word order preferences. I then report and discuss the results of Experiment 1.

1.1 A grammatical sketch of Basque

Basque is an OV language isolate with rich inflectional morphology, which allows both verb-medial and verb-final word orders. It is spoken in the Basque Country by about 700,000 people[4]. Basque presents a number of typological properties that are not found in other European languages: ergative alignment; rich agglutinative morphology; overt case on NPs; subject, object and dative verb agreement. Additionally, subjects, objects and datives can all be freely omitted (it is a three-way *pro-drop* language) (De Rijk, 1969, 2007; Hualde and Ortiz de Urbina, 2003). Given the agglutinative nature of its morphology, agreement morphemes are fully decomposable: verbal inflection agrees with the object (O), the subject (S) and the indirect object (IO) if there is one (see example 2). These agreement morphemes indicate whether these phrases are first, second or third person, singular or plural. Verbal auxiliaries immediately follow the verb except in negative and emphatic sentences where the auxiliary precedes the lexical verb (Laka, 1994). Although basic word order is S(IO)OV (De Rijk, 2007; Elordieta, 2001), Basque grammar allows significant freedom in the order of phrases in a sentence. In particular, constituents can be placed either before or after the verb, as illustrated in the set of examples in 3, where all variants are truth-conditionally equivalent. Example 3a represents canonical word order, commonly used when the sentence denotes discourse-new information. Other phrase orders are also possible, as shown in 3b - 3e.

(2) a. Kazetari-a-k gizon-a-ri egunkari-a-ø ekarri d-i-o-ø
Journalist-Det-Erg$_{sg}$ man-Det-Dat$_{sg}$ newspaper-Det-Abs$_{sg}$ brought Aux.Abs$_{3sg}$-Dat$_{3sg}$-Erg$_{3sg}$
The journalist has brought the man the newspaper

(3) a. S-IO-O-V:
Kazetariak gizonari egunkaria ekarri dio
Journalist-the-Erg man-the-Dat newspaper-the-Abs brought AUX
The journalist has brought the man the newspaper

b. S-O-V-IO:
Kazetariak egunkaria ekarri dio gizonari
Journalist-the-Erg newspaper-the-Abs brought AUX man-the-Dat

c. S-O-IO-V:
Kazetariak egunkaria gizonari ekarri dio
Journalist-the-Erg newspaper-the-Abs man-the-Dat brought AUX

d. O-IO-V-S:
Egunkaria gizonari ekarri dio Kazetariak
Newspaper-the-Erg man-the-Dat brought AUX journalist-the-Erg

e. etc.

[4]There is a total of 751.500 Basque speakers in the Basque-speaking territory, comprising north-eastern Spain and south-western France (VI Sociolinguistic Survey, 2016).

1.2 Availability-based models

Although most research on availability-based production has focused on the effects of conceptual accessibility (Bock, 1982; Bock and Warren, 1985; V. S. Ferreira, 1996; V. S. Ferreira and Dell, 2000, among others), several researchers have proposed a general preference to place short phrases before long ones (Arnold et al., 2000; de Smedt, 1994, *Easy First* principle in MacDonald, 2013; Stallings and MacDonald, 2011; Stallings et al., 1998; Wasow, 1997a, 1997b). The idea is that sentences such as 4b are preferred over sentences such as 4a, because short phrases (*to Bill*) should be generally easier for speakers to retrieve from memory than longer phrases (*the book she had been searching for since last Christmas*) and hence can be assembled earlier (example taken from Hawkins, 2004a: 26). This preference to produce short phrases before long ones should hold regardless of the typological properties of the languages involved. Indeed, the tendency to shift long constituents over short ones – often known as *heavy-NP shift* in generative grammar[5] (henceforth, HNPS)– is widely attested in corpora in VO languages like English, German[6], Hungarian, Greek, Polish, Finnish, Rumanian and Russian (de Marneffe et al., 2012; Hawkins, 1994; Hoffmann, 1999; Kizach, 2012; Köhler, 1999; Lohse et al., 2004; Temperley, 2007; Uszkoreit, 1987; Uszkoreit et al., 1998; Wasow, 2002; Wiechmann and Lohmann, 2013; for an overview, see Jaeger and Norcliffe, 2009).

(4) a. I [$_{VP}$ gave [[$_{NP}$ **the book** she had been searching for since last Christmas] [$_{PP}$ **to Bill**]]
b. I [$_{VP}$ gave [$_{PP}$ **to Bill**] [$_{NP}$ **the book** she had been searching for since last Christmas]]

There is also evidence from language production experiments supporting a *short-before-long* preference in English (Arnold et al., 2000; Stallings and MacDonald, 2011; Stallings et al., 1998). Stallings (1998) conducted a study that comprised of 3 experiments. Forty experimental items were constructed, consisting of a subject and a verb, a noun phrase and a prepositional phrase. Length was manipulated by adding prenominal adjectives and prepositional phrases to the short NP. In Experiment 1, participants had to read sentence fragments on a computer screen and plan the complete sentence they had to produce with those fragments. They indicated by pressing a key whether they intended to produce a S-V-NP-PP or a S-V-PP-NP sequence. In Experiments 2 and 3, participants did not have to indicate which structure they were going to produce. In Experiment 3, production of the sentence was cued by the S-V- fragment reappearing in the middle of the screen. In all three experiments, participants produced shifted orders about four times more often in the long NP condition than in the short NP condition, and thus reduced dependency length by placing the long constituent after the short one.

[5]I use the term that originated in the tradition of transformational grammar (Ross, 1967).

[6]Some languages such as Afrikaans, Dutch, Frisian, German and their regional varieties are not harmonic with respect to headedness. In German, for example, embedded clauses are head-final, whereas main ones are head-initial. For a more detailed description of headedness in the OV Germanic group, see Haider, (in press).

However, not all languages have demonstrated a *short-before-long* preference. As mentioned earlier, in Japanese, Korean and Persian, speakers tend to favor *long-before-short* word orders. Yamashita and Chang (2001) used a variant of Stallings et al.'s design (1998). Twenty-four ditransitive and thirty ditranstive sentences were constructed, and the length of the NPs was manipulated. Their results confirmed that speakers of Japanese tend to shift long NPs before short ones. In transitive sentences, constituents were shifted 27.5% more often when the object was long then when short. In ditransitive sentences, there were two types of shift, external (to the beginning of the sentence, before the subject or topic) and internal (to an embedded position in the middle of the sentence, after the subject or topic and before the indrect object). Objects were were shifted both to the internal-position (37.8%) and to the external-position (10.63%). The indirect object was only shifted to the beginning of the sentence (12.63%). Dennison (2008) conducted an experiment in which Korean native and non-native bilinguals participated. The design was identical to Yamashita and Chang's (2001), but only transitive sentences were used. Korean native speakers and Korean-dominant bilinguals produced *long- before-short* structures (29.4% and 41.9% respectively). However, English-dominant and Equi-dominant bilinguals showed a smaller preference for the *long-before-short* preference (3.3% and 6.8% respectively) (see Chapter 4, for a more detailed overview of these results and their interpretation). A last experiment in Persian, another head-final language, conducted by Faghiri et al. (2014), demonstrated a *long-before-short* preference in this language, confirming previous corpus findings (Faghiri and Samvelian, 2014). This experiment was conceptually similar to the ones reported above. 20 ditransitive sentences were created by alternating seven different verbs, and the length of the object was manipulated. Sentence constituents were presented unordered to the participants, who had then to write them down.

Some authors have suggested that the effect of availability is modulated by language-specific characteristics (Chang, 2009; Wasow, 2002, 2013; Yamashita and Chang, 2001). Yamashita and Chang (2001), who found a *long-before-short* preference in Japanese, argued that word order preferences are affected by both conceptual saliency and ease of formulation. Although, according to these authors, long phrases are more complex to formulate than short phrases, they are conceptually more salient because the additional words make the semantic representation of the head noun richer and stronger. English has a very strict word order, such that form-related factors exert a stronger effect, and because short phrases are easier to formulate than longer phrases, they tend to be placed before longer ones. In contrast, Japanese is a scrambling language with null pronouns (Saito, 2004), and that allows speakers to produce salient arguments early. Therefore, according to Yamashita and Chang (2001), the freer the word order, the stronger the effects of conceptual saliency, and hence the tendency to place long constituents before short ones (but see Tanaka, Branigan, and Pickering, 2011).

Wasow (2002, 2013) suggests that the general preference to place short elements before long ones could compete with other constraints. That is, in VO languages the *short-before-long*

word order might arise not only from the preference to place short elements earlier, but also from the preference to avoid center embedding, which is more costly to process (Chomsky, 1965; Gibson, 2000; Miller and Chomsky, 1963, *inter alia*). For instance, 4a is harder to process than 4b, because the relative clause embedded within the direct object increases the distance between the verb and the indirect object, making the integration of the indirect object harder. Therefore, speakers may prepose the indirect object to reduce the processing cost. In strictly OV languages, however, these two constraints pull in different directions: a *short-before-long* linearization of constituents places short elements earlier in the sentence, but it also creates a center-embedded structure. In these cases, speakers should seek to avoid center-embedding by shifting long constituents to a sentence-initial position, leading to a *long-before-short* ordering. This hypothesis about the opposing word order pattern in VO and OV languages is fairly similar to Hawkins' *Performance-Grammar Correspondence Hypothesis*, which I discuss in detail in the next section. However, it still makes the prediction that flexible OV languages such as Basque should prefer *short-before-long* orderings, since they could comply with the preference to place short elements earlier and still avoid complex constituents from appearing in the middle of a clause. In fact, Hualde and Ortiz de Urbina, in *A Grammar of Basque*, argue that Basque presents the same *short-before-long* tendency as English and other VO languages, and that the 5a and 6a patterns are preferred over the 5b and 6b (2003: 459-460):

> "Heavy constituents other than clauses tend to occur to the right of the clause unless focalized. In the following sentences with direct and indirect object, where neither of them is focalized, topicalized or dislocated, the heaviest one will normally occur to the right of the lighter element.

(5) a. Neuk emango diot Joni [bileran eskatutako informazio guztia]
I.Erg give.Fut AUX Jon.Dat meeting.Loc requested.Rel information all.Det
I will give Jon the information that was requested in the meeting

b. Neuk emango diot [bileran eskatutako informazio guztia] Joni

(6) a. Neuk emango diet informea [atzoko bileran eskatu zidatenei]
I.Erg give.Fut AUX report yesterday.Loc meeting.Loc request AUX.DAT
I will give the report to those who requested it at the meeting yesterday

b. Neuk emango diet [atzoko bileran eskatu zidatenei] informea]

Theoretical research in syntax has relied mainly on one type of data, informal introspective judgements by expert linguists. This has been proved to be a very productive methodology. Nevertheless, it also faces a number of challenges regarding its reliability, especially with regard to controversial judgments (Gibson and Fedorenko, 2010) and in less studied languages such as, for example, Hebrew (Linzen and Oseki, 2015) – or Basque in our case. Sprouse and Almeida (2012) argue that the use of acceptability judgment methodology is sound, and that for those cases where it might fail, peer review processes filter out any possible errors. However, that might be a more accurate statement for English than for

languages where peer-review is less extensive. Different solutions have been proposed for this state of affairs. On one hand, the creation of an online database of existing judgments where linguist can anonimously provide feedback and comments, together with the implementation of a peer review system for those judgments (Linzen and Oseki, 2015). On the other, using experimental techniques that allow affordable and fast collection of data (Gibson and Fedorenko, 2010; Gibson, Piantadosi, and Fedorenko, 2010). My study can be seen as a step forward in the direction of this second proposal, since it contrasts Ortiz de Urbina's intuition with data collected from naïve native speakers of Basque.

1.3 Hawkins' Performance-Grammar Correspondence Hypothesis (PGCH)

The PGCH model (Hawkins, 1994, 2004b, 2014) comprises three general principles that predict word order preferences: Minimize Domains, Maximize On-line Processing and Minimize Forms. The first two principles are relevant to the current study. The first principle that determines word order preferences is the Minimize Domains principle (henceforth, MiD). Both phrasal length and the head-direction of the language (OV vs. VO) determine word order preferences. Word orders that require fewer computations to reach a parsing decision about all the syntactic and semantic relations between the constituents within minimal processing domains are ranked as more efficient and are hence predicted to be more frequently produced.

The calculation of the size of a processing domain provides us with a metric that makes precise predictions regarding ranked frequencies of competing structures. This metric consists of the sum of the number of immediate (or sister) constituents (ICs) divided by the sum of the number of non-immediate constituents (words). This ratio quantifies the number of words that need to be processed in order to identify the immediate constituents of a given phrase, and thus it represents the level of processing efficiency of a given linguistic sequence. The higher the ratio, the more efficient the so-called processing domain. Hawkins predicts that word orders with optimal ratios will be produced more frequently than non-optimal ones. In other words, those constructions were fewer words are needed to resolve a dependency are to be preferred in production across all languages. For instance, in the example sentences 7 and 8 in English and Japanese below (taken from Hawkins, 2004a: 26 and Yamashita and Chang, 2001: B52) , 7a and 8a represent canonical word orders. In contrast, 7b and 8b represent shifted word orders, where the direct object is moved after (7b) or before (8b) the indirect object. MiD predicts that in both 7 and 8, the shifted orders (b) will be preferred over the non-shifted orders (a), since fewer elements (words or terminal nodes, in bold in the examples) are necessary to resolve the dependency, whose length is thus minimized.

(7) a. I [$_{VP}$ **gave** [$_{NP}$ **the book she had been searching for since last Christmas**] [$_{PP}$ **to** Bill]]
ICs: 3, {V, NP, PP}
Number of words: 12 {gave, the, book, she, had, been, searching, for, since, last, Christmas, to}
MiD ratio: 3/12 = 25%

b. I [$_{VP}$ **gave** [$_{PP}$ **to Bill**] [$_{NP}$ **the** book she had been searching for since last Christmas]]
ICs: 3, {V, NP, PP}
Number of words: 4 {gave, to, Bill, the}
MiD ratio: 3/4 = 75%

(8) a. Masako-wa [$_{VP}$ [$_{NP}$ **otoko-ni**] [$_{NP}$ **sinbun-de syookai-sarete-ita okasi-o**] **todoketa**]]
Masako-Top [$_{VP}$ [$_{NP}$ man-Dat] [$_{NP}$ newspaper-in appeared cake-Acc] delivered.]
Masako delivered the cake [which was] introduced in the newspaper to the man
ICs: 3, {V, NP, NP}
Number of words: 5, {otokoni, sinbunde, syookaisareteita, okasio, todoketa}
***MiD ratio*: 3/5 = 60%**

b. Masako-wa [$_{VP}$ [$_{NP}$ **sinbun-de syookai-sarete-ita okasi-o**] [$_{NP}$ **otoko-ni**] **todoketa**] Masako-Top [$_{VP}$ [$_{NP}$ newspaper-in appeared cake-Acc] $_{NP}$ man-Dat] delivered.]
Masako delivered the cake [which was] introduced in the newspaper to the man
ICs: 3, {V, PP, NP}
Number of words: 3, {okasio, otokoni, todoketa}
***MiD ratio*: 3/3 = 100%**

Consistent with Hawkins' predictions, corpus and experimental studies have consistently found a *short-before-long* preference in VO languages (e.g., English), but a *long-before-short* preference in OV languages (e.g., small corpus studies in Japanese, Korean and Turkish) (Choi, 2007; Hakuta, 1981; Hawkins, 1994; Kondo and Yamashita, 2011). Three production experiments provide evidence for a *long-before-short* tendency in Korean, Japanese and Persian (Dennison, 2008; Faghiri et al., 2014; Yamashita and Chang, 2001), as reported in the previous section. Further support for the PGCH hypothesis comes from computational models showing that short dependencies increase parsing accuracy (Collins, 2003; Eisner and Smith, 2005; Sleator and Temperley, 1991; Temperley, 2008, among others). As discussed above, only the PGCH model straightforwardly predicts a *long-before-short* preference for Basque. The efficiency ratios for transitive and ditransitive sentences in Basque are shown below (see examples 9a - 9c and 10a - 10c). The type of sentences used for the calculations of the efficiency ratios is identical to the ones used in our materials (see Section 2.3). According to these ratios, in Basque, in a sentence with a long object phrase, shifted orders where the object appears before shorter constituents are optimally efficient. That is, adjacency of phrasal heads reduces the number of words that need to be processed to complete a dependency, yielding more efficient structures. Therefore, one should expect that, in Basque transitive and ditransitive sentences, shifted O-S and O-IO orders are produced more frequently when the O is long compared to when it is short. Following the way in which Hawkins calculates efficiency metrics (Hawkins, 1994), the verb-participle and the inflected auxiliary count as two separate words in our calculations, which are presented in the set of examples below (9a - 9c and 10a - 10c), where efficiency ratios are provided for the sequences relevant for the purposes

of our study. Recall that the higher the ratio of a given sentence in comparison to others, the more efficient its processing will be and hence the more frequently it will be produced.

The second principle that guides word order preferences in the PGCH model, the Maximize On-line Processing principle (henceforth, MaOP), is concerned with speed in communication. It seeks to provide the earliest possible access to linguistic representation by avoiding processing delays and ambiguities. The general assumption is that many syntactic and semantic properties are not assignable to positions lower in thematic and syntactic hierarchy positions independently from the higher positions (agent > recipient > patient; c-commanding > c-commanded) (Primus, 1999) and that these assignment constraints will influence processing preferences, such that word orders that allow these assignments with minimum delays are preferable. MaOP predicts a general preference for subjects to precede objects and agents to precede patients, because, in cases when both the agent and the patient are present (that is, not in the case of unaccusative predicates), a patient requires a co-occurring agent, on which it is thematically dependent, and a c-commanded element requires a c-commanding one, on which it is syntactically dependent. Therefore, no preference for either S-O-V or S-V-O is predicted based on MaOP. Also, MaOP predicts that when the dependent category (patient, O) precedes the independent one (agent, S), both positions will be as adjacent as possible, because any intervening elements (e.g., V) will cause a delay in thematic assignment. Note that processing dependencies between O and V are assumed to be symmetric: O is dependent on V because it receives its thematic role from V, and V is dependent on O for semantic and syntactic disambiguation (i. e., *John ran* vs. *ran the race* vs. *ran the water* vs. *ran the advertisement*, from Keenan, 1979, as cited in Hawkins, 2014). According to MaOP, word orders with non-shifted NPs (e.g., S-O; IO-O) are favored, because shifted NPs alter the preferred word orders for property assignment. Therefore, in Basque either S-V-O or S-O-V in the transitive sentences and either S-IO-O-V or S-IO-V-O in the ditransitive sentences are considered optimal. Importantly, MaOP also predicts that if NPs are shifted, verb-final word orders, where the verb does not intervene between the noun phrases, should be preferred over verb-medial ones. This is in order to avoid any extra delay in the assignment of properties caused by the intervening V. Therefore, with respect to the MaOP principle, in Basque, S-O-IO-V should be preferred over other possible orders such as S-O-V-IO or O-V-S-IO in the ditransitive sentences and O-S-V over O-V-S in the transitive ones (see the efficiency ratios provided in 9a - 9c and 10a - 10c).

(9) *Ditransitive sentences*

a. S-IO-O-V:

[$_S$ **Aitak**] [$_{IO}$ **irakasleari**] [$_O$ **gurasoek** **gogoko** **duten** **pastela**] [$_V$ **ekarri**
[$_S$ Father-Erg] [$_{IO}$ teacher-Dat] [$_O$ parents like AUX-that cake-Det] [$_V$ took
dio]
AUX$_{O.IO.S}$]

The father brought the teacher the cake that parents like

'S: 2/7 = 28.6%

Number of words: 7, {aitak, irakasleari, gurasoek, gogoko, duten, pastela, ekarri}
ICs: 2, {NP, VP}
VP: 3/6 = 50%
Number of words: {irakasleari, gurasoek, gogoko, duten, pastela, ekarri}.
ICs: 3, {NP, NP, VP}
***MiD ratio*: 39.3% MaOP: high**

b. S-O-IO-V:

[$_S$ **Aitak**] [$_O$ **gurasoek gogoko duten pastela**] [$_{IO}$ **irakasleari**] [$_V$ **ekarri**
[$_S$ Father-Erg] [$_O$ parents like AUX-that cake-Det] [$_{IO}$ teacher-Dat] [$_V$ took
dio]
AUX$_{O.IO.S}$]
The father brought the teacher the cake that parents like
'S: 2/7 = 28.6%
Number of words: 7, {aitak, irakasleari, gurasoek, gogoko, duten, pastela, ekarri}
ICs: 2, {NP, VP}
VP: 3/3 = 100%
Number of words: {irakasleari, pastela, ekarri}.
ICs: 3, {NP, NP, VP}
***MiD ratio*: 64.3% MaOP: high**

c. O-V-S-IO:

[$_O$ Gurasoek gogoko duten **pastela**] [$_V$ **ekarri dio**] [$_S$ **aitak**] [$_{IO}$
[$_O$ Parents like AUX-that cake-Det] [$_V$ took AUX$_{O.IO.S}$] [$_S$ father-Erg] [$_{IO}$
irakasleari]
teacher-Dat]
The father brought the teacher the cake that parents like
'S: 2/3 = 66.7%
Number of words: 3, {ekarri, dio, amak}
ICs: 2, {NP, VP}
VP: 3/5 = 60%
Number of words: {pastela, ekarri, dio, amak, irakasleari}.
ICs: 3, {NP, NP, VP}
***MiD ratio*: 63.4% MaOP: lowest**

(10) *Transitive sentences*

a. S-O-V:

[$_S$ **Andreek**] [$_O$ **geltokian trenaren zain egon den gaztea**] [$_V$
[$_S$ Women-Erg] [$_S$ station-in train-of wait be AUX-that young person-Det] [$_S$
agurtu dute]
greeted AUX$_{O.S}$]
The women greeted the youngster that had been waiting for the train at the station
'S: 2/8 = 25%
Number of words: {andreek, trenaren, zain, egon, den, gaztea, agurtu}
ICs: 2, {NP, VP}
VP: 2/2 = 100%

Number of words: {gaztea, agurtu}
ICs: 32 {NP, VP}
***MiD ratio*: 62.5% MaOP: high**

b. O-S-V:

[$_O$ Geltokian trenaren zain egon den **gaztea**] [$_S$ **andreek**] [$_V$ **agurtu**
[$_O$ Station-in train-of wait be AUX-that young person-Det] [$_S$ women-Erg] [$_V$ greeted
dute]
AUX$_{O.S}$]
The women greeted the youngster that had been waiting for the train at the station
'S: 2/2 = 100%
Number of words: {andreek, agurtu}
ICs: 2, {NP, VP}
VP: 2/3 = 66.7%
Number of words: {gaztea, andreek, agurtu}
ICs: 2, {NP, VP}
***MiD ratio*: 83.4% MaOP: lower**

c. O-V-S:

[$_O$ Geltokian trenaren zain egon den gaztea] [$_V$ agurtu dute] [$_S$
[$_O$ Station-in train-of wait be AUX-that young person-Det] [$_V$ greeted AUX$_{O.S}$] [$_S$
andreek]
women-Erg]
The women greeted the youngster that had been waiting for the train at the station
'S: 2/3 = 66.7%
Number of words: {agurtu, dute, andreek}
ICs: 2, {NP, VP}
VP: 2/2 = 100%
Number of words: {gaztea, agurtu}
ICs: 2, {NP, VP}
***MiD ratio*: 83.4% MaOP: lowest**

Hawkins (2014) notes that, apart from minimizing the distance between phrasal heads, there may be other ways of enhancing efficiency. The MaOP principle predicts that syntactic processing in VO languages will tend to rely on agreement because it provides early access to the argument structure before all NPs are accessed. On the other hand, in OV languages, the parser will tend to rely on case-marking because it provides early online information about syntactic case and thematic role before the verb. However, NPs in Basque can be omitted or may be ambiguous because of case syncretism (for example, absolutive plural and ergative singular both end in *-ak*, see Hualde and Ortiz de Urbina, 2003). Thus, case marking alone does not always allow one to straightforwardly determine sentence structure. Also, in non-rigid OV languages with pluripersonal agreement as Basque, the early encounter of the verbal complex (lexical verb + auxiliary) can compensate for this. Specifically, Pablos (2011) has argued for Basque that agreement information provided by the auxiliary acts as a cue that

assists the parser in predicting upcoming elements[7]. Thus for instance, in a sentence like 11 the inflected auxiliary (highlighted) provides information about the upcoming indirect object (*gizon-a-ri*,"to the man") and object (*egunkari-a*, "the newspaper") (example adapted from Pablos, 2011).

(11) Kazetari-a-k ez **d-i-o-ø** gizon-a-ri egunkari-a ekarri
Journalist-Det$_{sg}$-S Neg Aux$_{3sg.O-3sg.IO-3sg.S}$ man-Det$_{sg}$-IO newspaper-Det$_{sg}$.O brought
'The woman did not bring the man the newspaper'

Placing the inflected verb early in the sentence facilitates the identification and integration of grammatical functions and semantic roles of both previous and upcoming constituents, because agreement obligatorily signals the number and type of arguments the sentence contains. I argue that when speakers of Basque produce constituent structures that are less preferable according to the MaOP principle (e.g., O-S over S-O, and O-IO over IO-O), placing the lexical verb and the agreement carrying auxiliary between the two constituents facilitates rather than delays the assignment of syntactic and thematic properties and thus this verb-medial order is preferred over immediate adjacency of shifted constituents. This is because less linguistic material needs to be accessed to determine initial sentence structure than in the case of shifted verb-final structures. For instance, in a *long-before-short* verb-medial ditransitive sentence in Basque (long-NP verb+auxiliary short-NP short-NP), when reaching the verbal complex, the previous absolutive-marked long constituent can be unambiguously identified as an object. Furthermore, the presence of the subject and the indirect object will be clearly indicated by verb-agreement, even before encountering the subject and indirect object phrases. In a *long-before-short* verb-final ditransitive sentence, sentence structure cannot be fully resolved until all constituents of the sentence have been accessed. One possibility is to assume that if different word orders have approximately the same MiD ratio, then the word order with a better MaOP ratio should be preferred (see Hawkins, 2004a for a discussion on the possible interplay of the three principles he proposes to account for cross-linguistic variation in grammars of the world). Note that non-rigid SOV languages with pluripersonal agreement can access initial structure either by case-marking (while shifted constituents remain adjacent to each other, consistent to MaOP predictions discussed earlier) or by placing the verb in a medial position (so that verb-argument structure is assigned immediately, by virtue of the agreement markers). Hawkins' proposal does not make any clear-cut prediction

[7]The semantic information carried by the lexical verb is often not sufficient to determine argument structure: for instance, the verb *hil* can mean either *die* when it is construed with one argument, or it can mean *kill* when constructed with two arguments, as shown in example 1 below:

(1) a. *die*: Katua hil da
Cat$_{Det-Abs.3sg}$ died Aux
'The cat has died'

b. *kill*: (Txakurrak) katua hil du
Dog$_{Det-Erg.3sg}$ cat$_{Det-Abs.3sg}$ killed Aux
'(The dog) has killed the cat'

as to whether the early position of the verb should be favored as more efficient than case marking in languages of this type. Thus, in addition to whether Basque deploys the same *long-before-short* preference as do rigidly verb-final SOV languages, I also examine whether verb-medial structures are preferred over verb-final ones when noun phrases have been reordered due to their length.

2 Experiment 1

2.1 Overview of the experiment

I investigated the effect of constituent-length on sentence word order. Specifically, I looked at the likelihood that speakers of Basque would produce shifted orders (O-S; O-IO) when the direct object is long versus short. In order to do so, I carried out a cued recall production task, where participants were presented with unordered phrases on a computer screen and were asked to construct transitive and ditransitive sentences using these phrases. I manipulated the length of the phrases (short vs. long) by adding relative clauses to them.

Let me briefly revise the predictions put forth in Sections 1.2 and 1.3 above. Efficiency-based approaches predict that speakers of any type of OV language, either rigid or flexible, will favor *long-before-short* word orders, where heads are placed adjacent to each other and distance across co-dependents is maximally reduced. These orders are claimed to be computationally more efficient than their non-shifted counterparts, which yield greater dependency distances. Thus, I expect Basque speakers to show a larger preference to shift the object before other constituents when it is long than when it is short (remember than when all elements are short no significant improvement in efficiency ratios is expected by shifting them), which would result in Long-O preceding S (O-S) in transitive sentences and IO (O-IO) in ditransitive ones. In contrast, if, as predicted by availability-based models, short phrases are easier to retrieve from memory than long ones, I then expect Basque speakers to show a stronger preference to shift the object after other constituents when it is long than when it is short, which would result in Long-O following S (S-O) in transitive sentences and following IO (IO-O) in ditransitive ones. Finally, if agreement morphology plays a role in efficient processing, as suggested above, one would expect speakers to produce more verb-medial sentences in contexts where they also produced more shifted orders than in contexts where they produced non-shifted orders. Thus, this will occur in sentences containing a Long-O if considerations of efficiency are at play. In turn, if the predictions of availability-based models apply, this will occur in sentences containing a Long-S for transitives or a Long-IO for ditransitives.

2.2 Participants

25 native speakers of Basque (15 females, mean age: 19.39) participated in the experiment. They were all undergraduate students at the University of the Basque Country (UPV/EHU). Adult speakers of Basque are bilingual with Spanish or French. All participants in this study were native speakers of Basque, and had Spanish as their second language. Data from one participant were excluded from the analysis due to technical problems with the recorder.

Proficiency and language use scores were obtained through a questionnaire (adapted from Weber-Fox and Neville, 1996) filled out by the participants before the experiment (see Appendix 1). The self-assessed index in Table 2.1 is the average of the participants' responses to four domains (speech comprehension, speech production, reading, and writing). The scores for proficiency for both Basque and Spanish are on a 7 point scale, in which 7 represents native speaker level and 1 bad level of proficiency. The scores for language use are on a 7 point scale, in which 7 stands for daily use of only Basque and 1 for daily use of only Spanish. The mean in Table 2.2 is the average of the participants' responses to three domains (home, school, others) along different periods of their life (childhood, adolescence, adulthood). They were also asked in which language the felt more at ease: 24 participants answered that they felt more comfortable using Basque, and 1 reported that s/he felt equally comfortable in both languages. In sum, for all of them Basque was their first language and their preferred way of communication. Participants all signed an ethical consent form and they were compensated economically for their participation.

	Basque	**Spanish**
Listening	6.83 (.38)	6.00 (.72)
Speaking	6.79 (.41)	5.00 (.88)
Reading	6.71 (.46)	5.67 (.86)
Writing	6.42 (.72)	5.17(.76)
Mean	6.69 (.49)	5.46 (.80)

TABLE 2.1: **Self-rated proficiency level scores (SD in brackets) in Experiment 1**. *Self-assessed language proficiency is rated on a 7-point scale: 7 = Native-like level; 6 = High level; 5 = Medium-high level; 4 = Medium level; 3 = Medium-low level; 2 = Low level; 1 = No knowledge*

	Infancy (0-3 years)	Primary school (4-12 years)	Secondary school (12-18 years)	Adulthood (after 18 years)
Home	1.25 (.44)	1.17 (.48)	1.21 (.51)	1.21 (.51)
School		1.50 (.66)	1.71 (.55)	1.71 (.95)
Other		1.71 (.55)	1.83(.48)	2.04 (.46)
Equally comfortable in both languages: 1; *More comfortable with Basque:* 24				

TABLE 2.2: **Language history (SD in brackets) as reported by participants in Experiment 1**. *Self-assessed language use is rated on a 7-point scale:* 1 = Only Basque; 2 = Mostly Basque, a few times Spanish; 3 = Mostly Basque but Spanish at least 25% of the time; 4 = Basque and Spanish equally; 5 = Mostly Spanish but Basque at least 25% of the time; 6 = Mostly Spanish, a few times Basque; 7 = Only Spanish

2.3 Materials

Two sets of experimental sentences were created, following the design in Yamashita and Chang, 2001. For the first set, 24 transitive sentences were generated (see Appendix 2). Three different versions of each sentence were created. I attached a relative clause to (1) no constituent or (2) the subject or (3) the direct object. This yielded the three main conditions of the Length variable: (1) short subject and short direct object condition (All-Short); (2) long subject and short direct object condition (Long-S); (3) short subject and long direct object condition (Long-O) (see Table 2.3). In the transitive sentences, the relative clause had a mean length of 12.2 syllables (SD = 2.5).

Phrase length condition	Example
All-Short	Andreek gaztea agurtu dute Lady$_{Erg.3pl}$ young person$_{Abs.3sg}$ greeted AUX$_{Abs.3sg-Erg.3pl}$ *The ladies greeted the youngster*
Long-S	[Geltokian trenaren zain egon diren andreek] gaztea agurte dute [Station-in train-of wait be AUX-that lady$_{Erg.3pl}$] young person$_{Abs.3sg}$ greeted AUX$_{Abs.3sg-Erg.3pl}$ *[The ladies that were waiting for the train in the station] greeted the youngster*
Long-O	Andreek [trenaren zain egon den gaztea] agurtu dute Lady$_{Erg.3pl}$ [station-in train-of wait be AUX-that young person$_{Abs.3sg}$] greeted AUX$_{Abs.3sg-Erg.3pl}$ *The ladies greeted [the youngster that was waiting for the train in the station]*

TABLE 2.3: **Example set of experimental transitive sentences used in Experiment 1**

For the second set, 30 ditransitive sentences were created (see Appendix 3). As before, I created three versions of each sentence, attaching a relative clause to (1) no constituent, (2) the indirect object or (3) the direct object. This yielded the three main conditions of the Length variable: (1) short direct object and short indirect object condition (All- Short); (2) short direct object and long indirect object condition (Long-IO); (3) long direct object condition and short indirect object condition (Long-O) (see Table 2.4). In the ditransitive sentences, the relative clause had a mean length of 10.5 (SD = 2).

To avoid effects due to reading habits, I manipulated the location on the screen of the NPs of the target sentences. For the transitive sentences, in one case the subject was placed in a box in the upper-right corner of the screen and the object was in the box in the bottom-left corner (S-O), whereas in the other case the positions were reversed (O-S). Similarly, for the ditransitive sentences, in one condition the indirect object was placed in a box in the bottom-left corner of the screen and the direct object was in the box in the bottom-right corner (IO-O), whereas in the other condition the positions were reversed (O-IO). In both transitive and ditransitive sentences the verb remained in the upper-left corner.

A norming study was run to test the compatibility of the relative clause with either the subject or the object in the transitive sentences and with either the indirect object or the direct object in the ditransitive sentences. Therefore, only the conditions in which the length had been manipulated were tested. Furthermore, only the conditions in which the components appear in their canonical order were subject to the norming study. The sentences were combined in two lists of 54 items each, so that each list would contain only one version of each sentence. Each list had 24 transitive sentences and 30 ditransitive sentences. There were no fillers included in the lists. They were then submitted to a normative study. 20 native speakers of Basque (12, first list; 8, second list) judged the acceptability of the sentences used on a scale of 1-5 (5 being totally acceptable and 1 being totally unacceptable). The mean rate was 4.8 and 4.4 when the S or IO were long and when the O was long respectively. In order to calculate the mean for both types of responses the data of four participants that completed the first list were randomly discarded in order to get the same number of responses as in the second list. All sentences received a rating of at least 4. Moreover, more than 50% of the sentences received a rating of 5. The instructions given with each list can be found respectively in Appendix 5.

2.4 Procedure

Six lists containing 54 experimental items (24 transitive and 30 ditransitive experimental items) and 56 fillers (see Appendix 2) were constructed. Across the six lists, all six versions of a given item set were presented only once. Each participant was presented with one list, in an individually randomized order, with the constraint that no two experimental sentences of the same set were presented next to each other. The experiment was programmed with DMDX

Phrase length condition	**Example**
All-Short	Amak irakasleari pastela eraman dio Mother$_{Erg.3sg}$ teacher$_{Dat.3sg}$ cake$_{Abs.3sg}$ take AUX$_{Abs.3sg-Dat.3sg-Erg.3sg}$ The mother brought the cake to the teacher
Long-IO	Amak [gurasoek gogoko duten irakasleari] pastela eraman dio Mother$_{Erg.3sg}$ [parents$_{Erg.3sg}$ like Aux-that teacher$_{Dat.3sg}$] cake$_{Abs.3sg}$ take AUX$_{Abs.3sg-Dat.3sg-Erg.3sg}$ The mother brought the cake [to the teacher that parents like]
Long-O	Amak irakasleari [gurasoek gogoko duten pastela] eraman dio Mother$_{Erg.3sg}$ teacher$_{Dat.3sg}$ [parents$_{Erg.3sg}$ like Aux-that cake$_{Abs.3sg}$] take AUX$_{Abs.3sg-Dat.3sg-Erg.3sg}$ The mother brought [the cake that parents like] to the teacher

TABLE 2.4: **Example set of experimental ditransitive sentences used in Experiment 1**

(Forster and Forster, 2003). Participants were tested individually. They first pressed the spacebar, which triggered the display of the first screen with four boxes containing sentence components. Participants prepared a sentence using the components given and pressed the spacebar when they were ready. After a blank screen for 500 milliseconds a simple arithmetic problem followed. The delay and the arithmetic problem were created in order to encourage participants to produce the sentences from their meaning, rather than by covert rehearsal (Yamashita and Chang, 2001). Participants pressed the spacebar when they solved the arithmetic problem. At this point, the verb of the sentence to be recalled was presented in the middle of the screen as a cue, and participants produced the sentence. Unlike in Yamashita and Chang (2001), the verb was presented in the middle of the screen, and it was not surrounded by empty boxes so as not bias any strategic responses (see Figure 2.1). Having produced the sentence orally, participants were prompted to respond to whether the sentence had appeared before in the experiment by pressing either Yes (key Z) or No (key M). Twenty of the 56 fillers were presented twice for this recognition task. The position of the verb in the fillers varied with respect to the experimental items. Also, the location of the arguments in the fillers was randomized. Experimental sessions were digitally recorded and transcribed for analysis.

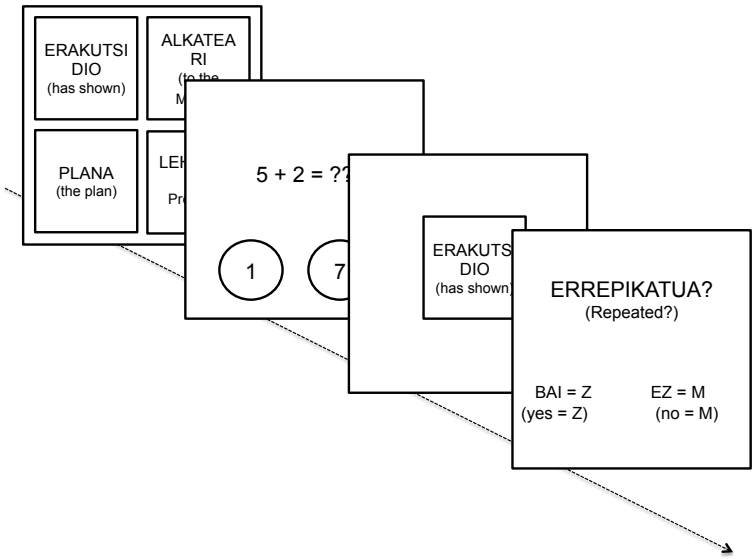


FIGURE 2.1: **Experimental procedure employed through Experiments 1 to 4**

2.5 Scoring

Participants' responses were scored in terms of (1) word order (shifted vs. non-shifted) and (2) verb position (medial vs. final). I scored *non-shifted order* when sentences were appropriately recalled and the constituents were arranged in their canonical word order (transitive sentences: S-O; ditransitive sentences: IO-O). I scored *shifted order* when sentences were appropriately recalled and the constituents were arranged in a non-canonical word order (transitive sentences: O-S; ditransitive sentences: O-IO). The position of the verb was coded as *medial* when the verb appeared between other constituents and as *final* when it appeared at the end of the sentence. I scored as miscellaneous sentences partial or no recalls, ungrammatical sentences, or when participants swapped the original length, function or number of any of the constituents, used an indefinite NP or employed non-target structures (n = 27, 2.1%). These miscellaneous responses were excluded from further analysis. In order to avoid data loss, however, responses in which participants substituted a word with a related one were included in the main analyses (n = 125, 9.6%; e.g., *proposamena* – 'proposal' – instead of *dokumentua* – 'document'). Our scoring resulted in 1269 target responses, 563 (97.7%) in the transitive sentences and 703 (96.5%) in the ditransitive sentences.

2.6 Analysis

I analyzed the number of shifted responses (relative to non-shifted responses) for transitive and ditransitive responses separately and the number of verb-medial responses (relative to verb-final responses) for ditransitives only. This last analysis was not conducted for transitive sentences due to lack of data. Because our dependent variables were binary, I analyzed our data using logit mixed effects models (Baayen, Davidson, and Bates, 2008; Jaeger, 2008). Our statistical analysis can overcome some of the drawbacks of more traditional statistical tests, since looking at the data only by subjects or by items can lead to spurious results. In order to avoid the language-as-a-fixed-effect fallacy (H. H. Clark, 1973). Fixed effects models such as ANOVA or t-test report a subjects analysis (F1) and an item analysis (F2). They are thus not appropriate for analyzing hierarchical data structures, in which leitems are nested within higher observations (subjects). Mixed models, on the other hand, allow simultaneous inclusion of by-participant and by-item variation. Morever, analysis of binomial data like our dependent variables (e.g., whether participants produced a non-shifted or shifted order) with ANOVA might lead to spurious results because confidence intervals can extend beyond the interpretable values of the categorical data (Jaeger, 2008). For example, in 10 elicited answers in an experiment, 8 answers are responded correctly (1) and 2 answers incorrectly (0). The mean will be .8 and the variance .18. In an ANOVA analysis that would result in a 95% confidence interval that would range from .52 to 1.08 (.8 $\pm$275), yielding an interpretation of the outcome variable as an impossible proportion of correct answers, and thereby likely underestimating the probability mass over events that actually can occur (see Jaeger, 2008 for a more detailed explanation).

In all models I included Length as a fixed effect. I collapsed across Position of the constituents on the Screen. Importantly, when the inclusion of the Position on the Screen predictor (i.e., in the main model of ditransitive sentences) yielded a better fit with the data, the results' significance of our critical Long-O condition was not altered in any of the models (all ps < .02). I always used the maximum random effect structure justified by the data by using model comparisons (Baayen et al., 2008). All data were analyzed using lme4 package (D. M. Bates, Maechler, Bolker, and Walker, 2014) of the statistical software R (version 2.14.2: R Core Team, 2014).

2.7 Results

2.7.1 Transitive responses

Table 2.5 reports the frequency and mean percentages of shifted and non-shifted responses and miscellaneous responses, as well as verb-medial and verb-final responses (see also 2.2). Table 2.6 reports the results of the analysis, which are also represented in Figure 2.2. All

figures included in this dissertations were made using ggplot2 package (Wickham, 2009). In the model where we regressed the transitive responses, our dependent variable was the production of shifted vs. non-shifted word orders. The Length predictor was Helmert-coded to avoid collinearity effects (after Helmert contrasts, there was no indication of collinearity, rs < .02). The Long-S contrast of the model in Table 2.6 compared the mean of All-Short with the mean of Long-S, and the Long-O contrast compared the mean of Long-O with the mean of All-Short and of Long-S. By-participant and by-item random intercepts were included.

Type of response			*Length condition*		
Constituent order	**Verb position**	**Structure order**	**All-Short**	**Long-S**	**Long-O**
S-O (non-shifted)	Final	S-O-V	185 (96.4%)	162 (88.5%)	147 (78.7%)
	Medial	S-V-O	0 (0%)	11 (6%)	3 (1.6%)
Total S-O			185 (96.5%)	173 (94.5%)	151 (80-3%)
O-S (shifted)	Final	O-S-V	2 (1%)	4 (2.2%)	9 (4.8%)
	Medial	O-V-S	5 (2.6%)	6 (3.3%)	28 (14.9%)
Total O-S			7 (3.6%)	10 (5.5%)	37 (19.7%)
Miscellaneous			0	9	4

TABLE 2.5: **Raw data and percentages (in brackets) of shifted and non-shifted responses relative to verb position in transitive sentences in Experiment 1**

	Estimate	**SE**	**z-value**	**p-value**
Transitive sentences (N = 563)				
Intercept	-4.147	.568	-7.301	<.001***
Length: Long-S	.590	.587	1.004	.316
Length: Long-O	2.350	.395	5.948	<.001***

TABLE 2.6: **Logit mixed model analysis for transitive sentences in Experiment 1**

The results of the main analysis revealed that participants produced more non-shifted orders (S-O) than shifted orders (O-S) when both constituents were short. Also, participants produced a similar number of non-shifted responses (S-O) in both the Long-S and the All-Short conditions. Importantly, speakers of Basque were more likely to produce the long object earlier in the sentence compared to a short constituent of the same type (All-Short, Long-S), which resulted in more shifted orders (O-S). However, I acknowledge that the high log-odds in the coefficients indicate that the model might be overfitted due to the relatively limited number of observations per cell (Jaeger, 2011), so our results should be interpreted with caution.

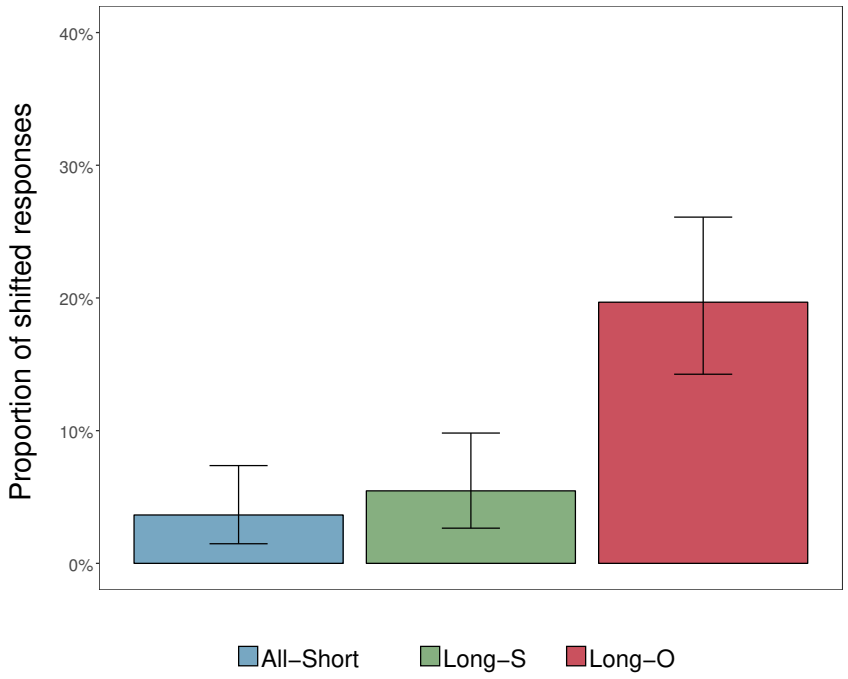


FIGURE 2.2: **Proportion of shifted responses in transitive sentences in Basque**. The error bars show 95% confidence intervals

2.7.2 Ditransitive responses

Frequency counts and means are reported in Table 2.7 (see also Figure 2.3). Note that all sentence word orders with less than 5% of occurrences in all conditions are grouped under Others. Although there were some miscellaneous responses (16, 2.2%), they were equally distributed across Length conditions ($p > .4$). Length was Helmert-coded to reduce collinearity [after Helmert contrasts, high collinearity levels (rs >.9) were reduced to mild-collinearity levels (rs. > .6)]. The Long-IO contrast of the main model in Table 2.8 compared the mean of All-Short with the mean of Long-IO, and the second contrast compared the mean of Long-O with the mean of All- Short and of Long-IO. By-participant and by-item random intercepts and a by-participant random slopes for Length were included as random effects. Table 2.8 reports the results of the analysis.

The results of the main analysis revealed that participants produced shifted orders (O-IO) and non-shifted orders (IO-O) similarly often when both constituents were short. Also, participants produced fewer shifted orders (O-IO) in the Long-IO condition than in the All-Short condition, but this difference did not reach significance. That is, they were more likely to produce the IO phrase earlier in the sentence when the IO phrase was long (Long-IO) than when it was not (All-Short). Importantly, participants were more likely to produce the long object (Long-O) earlier in the sentence compared to a short constituent of the same type, which yielded more *long-before-short* shifted orders (O-IO).

	Type of response		*Length condition*		
Constituent order	**Verb position**	**Structure order**	**All-Short**	**Long-IO**	**Long-O**
IO-O (non-shifted)	Final	S-IO-O-V	130 (54.9%)	117 (50.2%)	87 (37.3%)
		Others	0 (0%)	5 (2.2%)	0 (0%)
	Medial	IO-O-V-S	1 (.4%)	17 (7.3%)	0 (0%)
		Others	6 (2.5%)	17 (7.3%)	6 (2.6%)
	Total Final		130 (54.9%)	122 (52.4%)	87 (37.3%)
	Total Medial		7 (2.9%)	35 (14.6%)	6 (2.6%)
Total IO-O			137 (57.8%)	157 (67%)	94 (39.9%)
O-IO (shifted)	Final	Others	7 (3%)	3 (1.3%)	11 (4.7%)
	Medial	S-O-V-IO	88 (37.1%)	71 (30.5%)	89 (38.2%)
		O-V-S-IO	4 (1.7%)	0 (0%)	36 (15.5%)
		Others	0 (.4%)	3 (1.2%)	4 (1.7%)
	Total Final		7 (3%)	3 (1.3%)	11 (4.7%)
	Total Medial		92 (39.2%)	74 (31.7%)	129 (55.4%)
Total O-IO			99 (42.2%)	77 (33%)	140 (60.1%)
Miscellaneous			4	4	6

TABLE 2.7: **Raw data and percentages (in brackets) of shifted and non-shifted structures relative to verb position in ditransitive sentences in Experiment 1** (verb-medial and verb-final orders with less than 5% occurrences in all conditions were subsumed under Others)

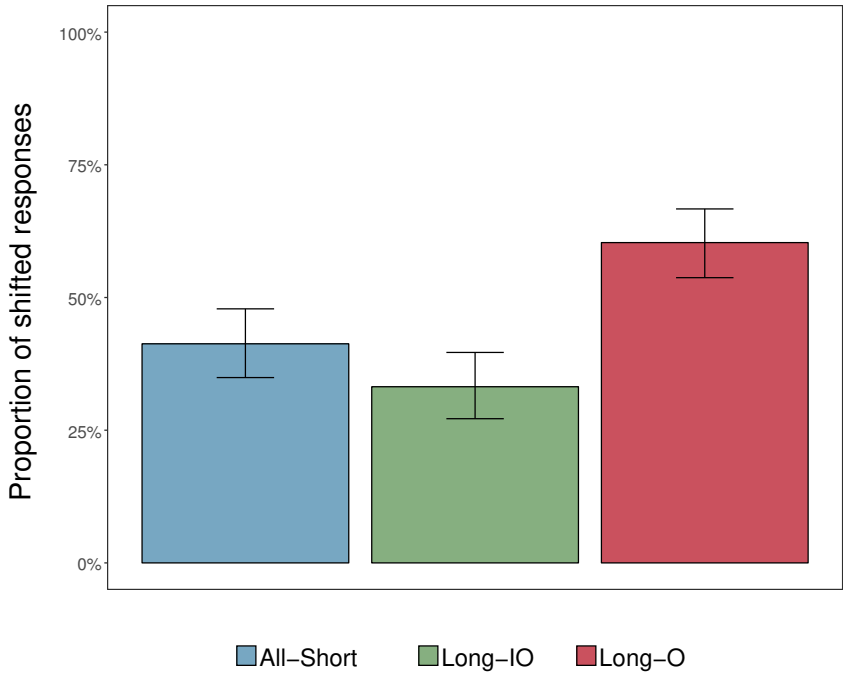


FIGURE 2.3: **Proportion of shifted responses in ditransitive sentences as a function of dependency length in Basque**. The error bars show 95% confidence intervals

	Estimate	**SE**	**z-value**	**p-value**
Ditransitive sentences (N = 703)				
Intercept	-.278	.321	-.866	<.386
Length: Long-IO	-.624	.317	-1.970	.049*
Length: Long-O	1.417	.525	2.700	.006**

TABLE 2.8: **Logit mixed model analysis for ditransitive sentences in Experiment 1**

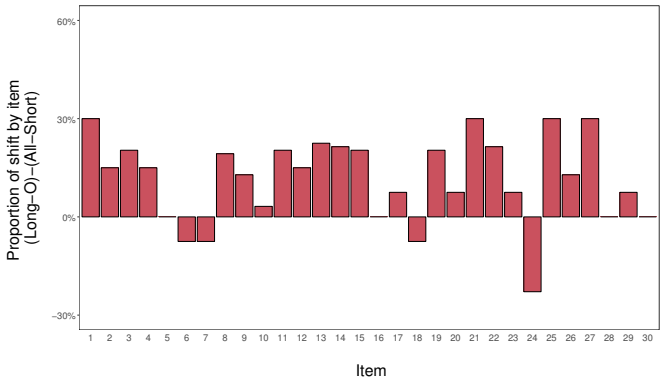


FIGURE 2.4: **Proportion of shifted responses by items in Experiment 1.** Negative proportions indicate a *short-before-long* ordering. No barplot indicates a preference for canonical orders. Positive proportions indicate a *long-before-short* ordering

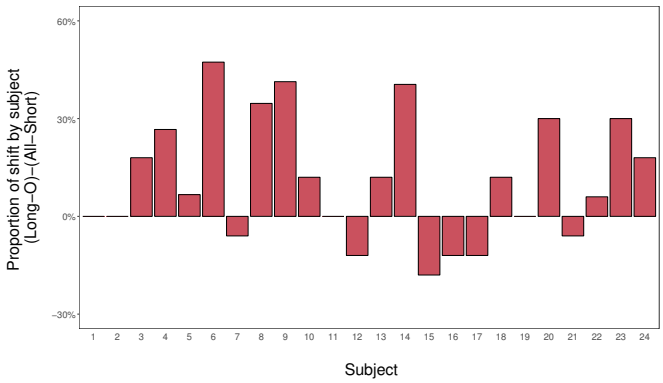


FIGURE 2.5: **Proportion of shifted responses by subjects in Experiment 1.** Negative proportions indicate a *short-before-long* ordering. No barplot indicates a preference for canonical orders. Positive proportions indicate a *long-before-short* ordering

To ensure that effects of length were not due to a small subset of the items or the subjects, the results for the individual ditransitive sentences were examined. Of the 30 sentences, 4

yielded no shift. Of the remaining 26, 22 yielded shifts that placed the long phrase before the short one (see Figure 2.4). Differences across items were also reported in Stallings et al. (1998), who found that argument structure biases impact production processes (for similar results in comprehension data, see MacDonald, 1994; MacDonald, Pearlmutter, and Seidenberg, 1994; Trueswell, Tanenhaus, and Kello, 1993). Similarly, of the total 24 subjects, 4 shifted no sentences when the object was long. Of the remaining 20, 14 showed a preference to shift long elements before short ones and 6 showed the opposite tendency to place long elements after short ones (see Figure 2.5). That is, a majority of our subjects and items showed a preference for *long-before-short* orders. Additionally, there were numerical differences across item and subjects with respect to the magnitude of the effect. I will address the differences across subjects in Chapters 4 and 5.

2.7.3 Verb position in ditransitive responses

I modeled the production of verb-medial and verb-final responses with the aim of exploring whether the production of shifted structures led to the production of verb-medial over verb-final sentences. Given that verb-medial responses were generally rare, I could not meaningfully analyze if verb position was conditional on word order shift. I thus indirectly examined this issue by analyzing the log-odds of verb-medial and verb-final responses as a function of Length in the ditransitive model, where I had enough responses. By-participant and by-item random intercepts and a by-participant random slope for Length were included as random effects. Because there was no sign of collinearity in the ditransitive model ($rs < .01$), Length was effect-coded. Results of the analysis are reported in Table 2.9.

The results of the verb-position analysis revealed a non-significant intercept indicating that participants produced verb-medial and verb-final responses similarly often. Participants did not produce more verb-medial sentences when the IO phrase was long (Long-IO) than when it was not (All-Short). In contrast, they did produce more verb-medial responses in sentences containing a long object (Long-O) compared to when they contained a short subject and a short object. That is, in those circumstances where the production of a non-canonical O-IO order was more likely (see Figure 2.6).

In sum, the results from this experiment reveal an overall tendency to use canonical orders, but only for transitive sentences. Crucially, when Basque speakers do resort to shifted word orders, they prefer to place long constituents before short ones, both in transitive and ditransitive sentences. Additionally, evidence from the production of ditransitive sentences showed that, in those contexts in which speakers showed a preference to produce shifted word orders (i.e., in Long-O conditions), they also showed a preference to place the verb earlier in the sentence (i.e., in a verb-medial position).

	Estimate	SE	z-value	p-value
Verb position (N = 703)				
Intercept	-.484	.351	-1.380	.168
Length: Long-IO	.178	.298	.597	.551
Length: Long-O	.909	.418	2.178	.029*

TABLE 2.9: **Logit mixed model analysis for verb position in ditransitive sentences in Experiment 1**

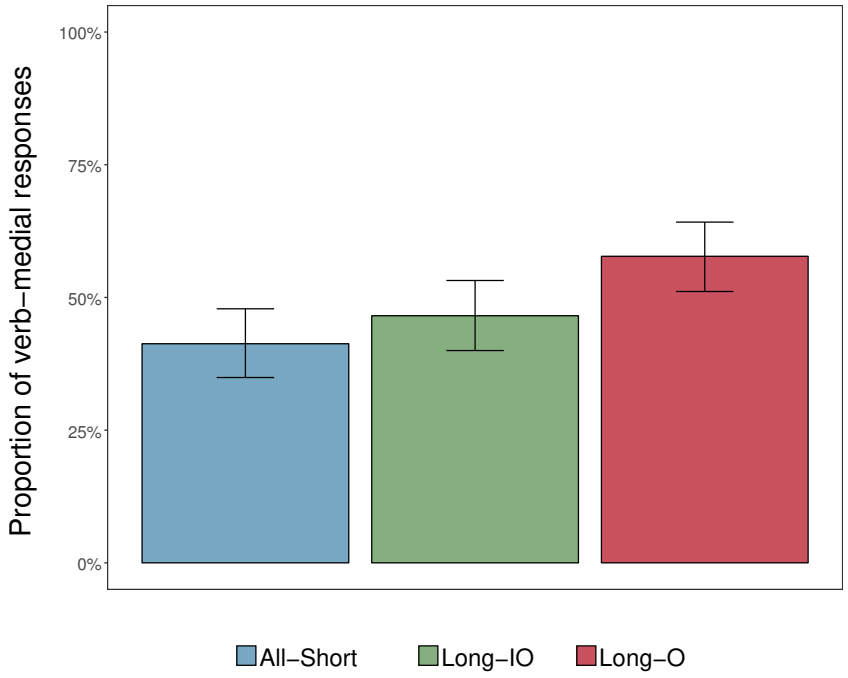


FIGURE 2.6: **Proportion of verb-medial responses ditransitive sentences as a function of dependency length in Basque**. The error bars show 95% confidence intervals

3 Discussion

In this chapter I examined the effect of phrasal length on sentence word order preferences in a flexible sentence word order OV language, Basque, a language that has both case-marking and pluripersonal verb agreement and allows for both verb-medial and verb-final orders. I found that the speakers of Basque fronted the object more frequently when it was long than when it was short, revealing a general preference to produce long phrases before short ones, as has been found in rigidly verb-final OV languages (Dennison, 2008; Yamashita and Chang, 2001). I can thus conclude that speakers of OV languages do not comply with a general *short-before-long* preference, contrary to what was predicted by availability-based models

(Arnold et al., 2000; *Easy First* principle in MacDonald, 2013; Stallings and MacDonald, 2011; Stallings et al., 1998; Wasow, 1997a, 1997b) and to what was argued for Basque in some grammatical descriptions and normative style books[8]. Our findings support the prediction made by Hawkins' PGCH that speakers of OV languages will favor *long-before-short* sentence word orders (Hawkins, 1994, 2004a, 2014), and that the *short-before-long* preference is exclusive to VO languages. That is, typologically different languages resort to mirror-like orders that comply with the general processing principle of minimizing dependencies and so yield minimal computational domains, in Hawkins' terms. Note that this should not be taken to mean that availability effects do not play any role in OV languages. Rather, in our data dependency minimization strategies seem to play a more important role during sentence production planning (e.g., there is plenty of evidence showing availability effects in OV languages: Branigan et al., 2008; V. S. Ferreira and Yoshita, 2003, *inter alia*). Some of the findings reported in this chapter are not straightforwardly predicted by the PGCH or other related approaches, and I will discuss them in the remainder of this chapter.

3.1 Verb-medial responses

Our results reveal a preference for verb-medial sentences in ditransitive sentences when the object is long and has been shifted. I argued that speakers produce the verb early to compensate for processing difficulties. As discussed earlier, verb agreement directly signals thematic assignment in Basque and can guide the exact prediction of upcoming sentence structure, so that when the verb is fronted, agreement morphology provides enough information to fully determine sentence structure (see Pablos (2011) for evidence for this claim). Therefore, the position of the verb is likely to affect processing efficiency in Basque, thus impacting word order preferences.

An alternative explanation for the preference for verb-medial sentences in Basque might be similarity-based interference (MacDonald, 2013). It has been shown that when semantically related nouns are planned and uttered close to each other, production difficulty increases (M. Smith and Wheeldon, 2004), possibly because speakers need to inhibit one of the nouns when producing the other but later have to retrieve it. MacDonald and colleagues argue that this similarity-based interference may influence the choice of utterance form (Gennari et al.,

[8]In the examples provided by Hualde and Ortiz de Urbina in support for their claim that long or heavy elements tend to appear to the right of short ones Hualde and Ortiz de Urbina, 2003: 459-460), cited in Section 1.2, the two elements whose relative ordering is relevant for their claim, the direct and indirect object, appear in a postverbal position, with a relative clause attached to the direct object. In this specific case, either the *short-before-long* or the *long-before-short* ordering of the two phrases inside the VP is equally efficient. In both sentences 5a and 5b, 7 words need to be parsed to identify 3 immediate constituents (VP, NP, NP), yielding an efficiency ratio of 23.3% for each of them. In this case, according to Hawkins' claims, these two sentences should be used with a similar frequency, since their efficiency ratios are identical. Also, the effect of phrasal length needs to be considered when no other factor, such as discourse considerations, is also at play. In these example the subject is focalized: the form *neuk* and not the unmarked *nik* is used. In Basque, focus is generally preverbal and adjacent to the verb (Altube, 1929; Ortiz de Urbina, 1989).

2012; MacDonald, 2013). Speakers produce passive subject relatives more than active object relatives when two animate nouns are to be mentioned than when one noun is animate, and the other inanimate (Gennari et al., 2012). One may thus wonder whether Basque speakers tend to produce verb-medial responses more when they have to produce similar animate NPs adjacently. However, although an interference-based account may explain why speakers sometimes produced transitive OVS sentences ($NP_{animate}$-V-$NP_{animate}$) in Basque, it is unclear how it can explain verb-medial responses for ditransitive sentences, where two animate NPs appear at the same side of the verb (OVSIO: $NP_{inanimate}$-V-$NP_{animate}$-$NP_{animate}$). Most importantly, this account does not straightforwardly explain why verb-medial responses were affected by the length of the constituents (i.e., in the case of transitive sentences no verb-medial responses were produced when all the NPs were short in All-Short conditions regardless of animacy of the noun phrase).

Furthermore, the *Noisy-Channel account* (Gibson et al., 2013) may explain why Basque speakers sometimes produced SVO utterances. According to this account, SVO arises as a result of communicative pressures, due to users' sensitivity to the possibility of noise corrupting the linguistic signal. When a potential ambiguity arises in an SOV word order (e.g., in the case of semantically reversible events, where two possible agents precede the verb), a shift to SVO is predicted in order to maximise the listener's ability to recover the meaning, given the noisy process. Therefore, this account predicts that, in semantically reversible sentences like the transitive ones in our study, where case marking is ambiguous (an absolutive singular NP can be an intransitive S or an O), the number of SOV word order sentences will decrease and SVO orders will increase. However, our participants produced SVO orders only very rarely. In fact, they produced SVO word orders 6% of the time when the subject was long and 1.6% when the object was long. No SVO orders were produced when all constituents were short. Also, there was an overall preference for canonical SOV word orders in transitive sentences in Basque, consistent with corpus studies (De Rijk, 1969; Pastor and Laka, 2013). It is also possible that the hypothesized shift from SOV to SVO is restricted to situations where case marking is completely ambiguous and inefficient to disambiguate the interpretation of reversible events, and our subjects in transitive sentences were plural marked, rendering case-marking fully unambiguous.

3.2 Differences across languages

The effect of length on word order appears to be smaller in Basque than in other languages in similar experimental conditions. The rate of shifting in transitive sentences was only 16.1% in Basque, whereas it reached 26% and 28.2% in Japanese and Korean respectively (Dennison, 2008; Yamashita and Chang, 2001). In ditransitive sentences (for which no Korean data exist), Japanese also resorted to shifting more often (40.3%) than Basque (17.9%). In a study on English involving long and short PPs, Stallings and MacDonald (2011) reported that length affected word order in about 40% of cases. In any case, in this last study, different materials and types

of structures were used, so these differences need to be treated cautiously. Here I provide a series of tentative approaches to account for this difference, and address this issue fully in the next chapter.

The relative strength of processing preferences depends on the availability of cues that signal grammatical structure (e.g., word order, case marking, agreement) (MacWhinney, 1987). If more than one cue is available in a given language, the most frequent and hence reliable cue should have the largest impact on processing strategies. This hypothesis has been supported by data from English, German and Russian (Kempe and MacWhinney, 1999; MacWhinney and Bates, 1989; for related claims see Kurumada and Jaeger, 2015; Temperley and Gidea, 2010) and by data from an artificial language learning study (Fedzechkina et al., 2012, 2013). Hence, it is possible that English, Japanese and Korean speakers shifted word order more often than Basque speakers because word order is a significantly less reliable processing cue in Basque than in the other languages. Also, other researches have suggested that rich morphological systems might interact with the preference for minimizing dependency length (Futrell et al., 2015b; Gildea and Jaeger, 2015; Gildea and Temperley, 2010; Tily, 2010), by reducing the possibility of interference during retrieval or encoding of distant co-dependents. Another possibility is that word order-based expectations, which might be different across languages, interact with the preference towards reducing the distance between co-dependents. All these approaches are broadly compatible with information-theoretic and inference-based accounts that link production preferences to efficient communication (Gibson et al., 2013; Jaeger, 2010; for further discussion and references, see Jaeger and Tily, 2010). I will tackle these two last possibilities in Chapter 4. Finally, it is plausible that, as other means such as case marking are used to signal underlying syntactic structure, word order is recruited to other purposes such as conveying information structure and therefore the impact of constituent length on sentence word order might lessen as a result (Gildea and Temperley, 2010).

Yamashita and Chang (2001) have also related word order freedom and production strategies, although in a different manner. They argue that long phrases are conceptually more salient, and speakers will tend to produce conceptually salient information early, to the extent that it does not conflict with other word order constraints such as fixed word order. Specifically, they argue that the freer the word order of a language, the stronger the effects of conceptual saliency and hence the tendency to place long elements before short ones. My results do not support this hypothesis. As discussed before, Basque seems to show a smaller degree of HNPS as a result of phrase length than Japanese or Korean, even though it has a freer sentence word order. In fact, Futrell (2017) reports that, in a sample of 37 languages, the greater the word order freedom of a language, the weaker its reliance on the preference to minimize dependency lengths or, in Hawkins' terms, processing domains. Thus, conceptual saliency does not appear to be a critical factor in cross-linguistic word order differences relative to dependency length.

4 Concluding remarks

Our study sheds new light on the factors determining sentence word order preferences, in particular on the preference for dependency length minimization across languages, and it contributes to our understanding of how language processing is modulated by demands interacting with variations in the properties of grammars. The Performance-Grammar Correspondence Hypothesis (PGCH) (Hawkins, 1994, 2004a, 2014) and related accounts (Futrell et al., 2015b; Gildea and Temperley, 2010; Temperley, 2008, among others) claim that minimizing dependency distance makes language processing more efficient. The PGCH, in particular, predicts a *short-before-long* preference in VO languages and a *long-before-short* tendency in OV languages. This general prediction regarding the different preferences in VO versus OV languages is met in this study, where a non-rigid OV language was investigated. Specifically, when the sentence word order produced by our participants departed from the canonical one, long noun phrases preceded short ones, contrary to predictions from availability models. Additionally, the impact of dependency length minimization seems weaker in Basque than in other languages, and variable across subjects. I address this issue in following chapters.

Hawkins (1994, 2004a, 2014) proposes a metric that measures the degree of processing efficiency, yielding precisely ranked predictions regarding sentence word order preferences in language production. This metric predicts that *long-before-short* verb-final sentences will be maximally efficient in Basque and thus will be preferred by speakers. These predictions were not fully consistent with the results reported in this study. When the object was long, and the production of shifted orders is thus more likely, verb-medial sequences were favored over their verb-final counterparts. I argue that verb agreement is the factor behind this preference for shifted verb-medial word orders. In contrast to rigidly verb-final OV languages that lack verb agreement, in heavily inflected flexible OV languages like Basque, a fronted verb that carries agreement morphology is a very reliable cue to sentence structure. In other words, differences in the morpho-syntactic cues available in the language yield different patterns of sentence word order preferences, although the principles behind them are common to all languages. All in all, this first chapter clearly shows that cross-linguistic research is crucial for understanding how universal processing demands interact with different types of grammatical specifications, yielding different production patterns that emerge from a common source.

5 Summary of findings

1. Unlike what is claimed in some descriptive and prescriptive works in Basque, and contrary to predictions from availability-based models, Basque, a non-rigid head-final language, favors *long-before-short* and not *short-before-long* orders. This constitutes

additional evidence in favor of the claim that the bias for dependency length minimization is a universal processing principle.

2. Language-specific properties such as verb agreement impact the type of mechanisms by which dependency length is reduced cross-linguistically. Basque has a preference for verb-medial orders when the direct object is long, which constitutes evidence for an active role of agreement morphology in enhancing efficient processing by signaling the thematic roles of upcoming verbal arguments.

Chapter 3

Cross-linguistic differences in the bias for short dependencies

Abstract

Language form is claimed to be shaped by a cognitive bias for short dependencies due to universal constraints on working memory and/or general requirements on efficient communication. However, large-scale cross-linguistic corpus studies find residual variance between languages with regard to this bias. Also, in Chapter 2 Basque showed a weak preference for short dependencies. In this chapter, I tested whether there are cross-linguistic differences in the magnitude of the universal preference for dependency length minimization and whether these differences are determined either by case morphology or by word order-derived expectations. Rich case marking morphology provides cues that add memory retrieval and reduce uncertainty when resolving long dependencies. Alternatively, certain word orders could reduce the level of predictability and processing cost of co-occurring linguistic items. Both possibilities would translate into a weaker reliance on short dependencies. I compared the effect of dependency length on sentence word order preferences in three languages (Basque, Polish and Spanish) that differ with respect to the richness of their case morphology (Basque and Polish versus Spanish) and basic word order (Basque - OV - versus Polish and Spanish - VO). The results from Experiment 2 indicate that there is very strong variation in the preference for short dependencies *across* languages and that this variation is not random. Head-initial Polish and Spanish pattern identically in transitive and ditransitive sentences, and radically differ from head-final Basque. In ditransitive sentences, Polish and Spanish show a much stronger tendency to reduce dependency length than head-final Basque. In transitive sentences, Polish and Spanish present no effect of dependency length, whereas Basque does. This variation can be derived from differences in language-specific word order-based expectations of upcoming information, and not from differences in richness of case morphology. I argue that knowledge of language-specific distributions could limit the general preference for reducing dependency

length and that further cross-linguistic research on the upper and lower bounds on efficient processing and on its interaction with other factors such as discourse considerations is needed.

RESEARCH QUESTIONS:

- Does the magnitude of the bias for short dependencies vary across languages?
- If it does, do differences in richness of case morphology affect the strength of the preference for short dependencies?
- If it does, do word order based-expectations affect the strength of the preference for short dependencies?

1 Introduction

Recent evidence indicates that languages tend to shorten the distance between elements in a dependency. Large-scale corpus studies provide cross-linguistic evidence for an above-chance preference for short dependencies in up to 37 languages (Futrell et al., 2015b; Gildea and Jaeger, 2015; Liu, 2008), confirming findings from smaller corpora (Gildea and Temperley, 2010; Hawkins, 1994). The tendency to reduce dependency length has been claimed to modulate the frequency of word orders within and across languages (Hawkins, 1994), consistency (Greenberg, 1963; Hawkins, 1983 but see Dunn et al., 2011) and inconsistency in head direction (Dryer, 1992), the form of possible grammars (Futrell et al., 2015b; Liu et al., 2017, *inter alia*) and language change (Liu and Xu, 2012; Tily, 2010). This preference for short dependencies attested in corpora has been linked to robust findings in comprehension, where there is ample evidence that the distance across co-dependents increases processing difficulty during integration processes (Bartek et al., 2011; Gibson, 1998, 2000; Grodner and Gibson, 2005), possibly due to degradation of the memory mechanisms involved (Lewis and Vasishth, 2005). The tight convergence between production and comprehension data has been interpreted as evidence that human languages follow a general principle of rational action that facilitates processing and promotes efficient communication through different mechanisms such as dependency length minimization (DLM) (Fedzechkina et al., 2018).

The preference for short dependencies is described as a property of all languages derived from limitations in the memory resources needed during language processing (Liu et al., 2017) and from a general bias for efficient communication (Fedzechkina et al., 2018; Gildea and Jaeger, 2015; Hawkins, 1994). Surprisingly, however, some corpus studies suggest that there might be cross-linguistic differences in the strength of the preference for short dependencies, which, unless accounted for, would suggest quite variable upper and lower thresholds for either processing cost or communicative efficiency across languages (see Liu, 2008 for the claim that Chinese exerts a stronger taxation on working memory than other languages). To account for these differences it has been speculated that either head-finalness or rich

morphology could impact the tendency to minimize dependency length (Futrell et al., 2015b; Gildea and Jaeger, 2015; Gildea and Temperley, 2010). It has been suggested that case morphology could reduce the risk of ambiguity when co-dependents are distant from one another (Gildea and Temperley, 2010). Also, rich word-internal structure (derivational morphology) could encode information that less inflected languages convey by means of word order (Gildea and Jaeger, 2015). Gildea and Temperley (2010) found a stronger preference for short dependencies than expected by chance in English compared to German (see also Park and Levy (2009) for similar results), a "partially" head-final language with a richer morphological system than English[1]. Additionally, in a sample of five languages (English, Arabic, Mandarin, Czech and German) investigated in Gildea and Jaeger (2015), German was the only one that presented non-optimal processing efficiency based on depenedency length and local surprisal alone.

As stated above, basic word order and case marking are highly correlated: head-final languages tend to have rich case morphology (Blake, 1994; Greenberg, 1963). Using a sample of 502 languages, Dryer (2002) calculates that a 72% of head-final versus 22% of head-initial languages (counting both VSO and SVO) have some form of case marking that distinguishes the two nominal arguments in a transitive clause. Rich case morphology, head-position and word order freedom are highly correlated typological traits (Dryer, 2002). In all the studies mentioned above, where cross-linguistic differences where never explicitly tested (except in Gildea and Temperley, 2010), it is thus unclear whether any differences across languages could be due to either case morphology, head position, word order freedom or all of them. Furthermore, other studies present contradictory evidence with regards to the claims that either morphology or head position affect the preference for dependency length minimization. Across two samples of 22 and 20 typologically diverse languages respectively, Liu (2008) and Liu and Xu (2012) found that Chinese, an analytic language with virtually no inflectional morphology, shows the weakest tendency to minimize dependency length. More crucially, no correlation between a bias for reducing dependency length and head position was found in this study. Also, in Liu (2008), two head-final languages such as Japanese and Turkish show smaller dependency length means than head-initial and morphologically poorer languages such as Chinese or Spanish. In any case, so far, the role of head position and case marking morphology on the strength of the preference for short dependencies has not been systematically studied.

In this study, for the first time as far as I know, I investigated whether the cognitive bias for short dependencies varies across languages by controlling crucial grammatical properties that tended to be confounded in the available data so far, namely case marking and basic word order. In order to explore the bias for shortening dependencies in these three languages, three languages were compared, Basque, Polish and Spanish, that differ with respect to case marking or lack thereof (rich in Basque and Polish, poor in Spanish) and word order (head-final Basque versus head-initial Polish and Spanish). Specifically, a controlled production experiment was

[1]See foonote 4 in Chapter 2

carried out, where speakers had the choice of alternative word orders to express the same proposition (see Sections 2.4 in Chapter 2 and 2.4 below in this chapter). In all three languages, transitive and ditransitive constructions with similar meanings and lengths were used.

Experimental data can greatly contribute to recent findings emerging from large-scale corpus studies. On the one hand, corpus studies are mainly based on written data, which opens up the possibility that any preference for short dependencies is limited to texts specifically edited for reading (but see Gildea and Jaeger (2015) and Futrell et al. (2015b) for an above-chance measure of dependency length minimization of spoken data in English and Japanese respectively). Also, as Hudson (2017) points out, there might be different requirements on the degree of textual complexity allowed across cultural communities. On the other, the genres of texts used for each language need to be carefully selected. For instance, by reducing the number of poetic texts included in their corpus, Gulordava and Merlo (2015) found that, in contrast to Futrell et al. (2015b), Latin and Ancient Greek (head-final) show a strong tendency for short dependencies. A small effect of genre on dependency length as also been reported for English (Wang and Liu, 2017). Additionally, cross-linguistic differences might be spuriously derived from mixing dependency lengths out of sentences of different lengths (Ferrer i Cancho and Liu, 2014) or from comparing actual dependencies with random baselines that are not tolerated by the grammar of the language(s) (Ferrer i Cancho and Liu, 2014; Liu, 2008). Finally, especially relevant for cross-linguistic comparisons is the lack of an appropriate parallelism in labelling grammatical relations expressed by prepositions versus morphology (Zeman et al., 2012). For example, prepositions and clitics are counted as independent words in Spanish, whereas in Basque or Polish postpositions and agreement morphemes are not (Aduriz et al., 2003; Taulé, Martı, and Recasens, 2008). Differences in annotation schemes can thus lead to an overestimation of the proportion of short dependencies in case marking languages such as Basque or German as opposed to languages as English (see set of sentences 12a - 15b below, where orthographic word count was used as a fitting proxy for dependency length, as is customarily done in corpus studies, and the resulting dependency length for Basque or Polish is inevitably smaller than for Spanish). Finally, as already mentioned in Chapter 1, there were relevant differences regarding the baselines employed in these large-scale corpus studies (see 1.2 for a more detailed explanation), that could have impacted their result

All these issues were addressed by conducting a production study in three languages that differed in case marking (rich/poor) and basic word order (OV/VO), namely Basque, Spanish and Polish. In Section 1.1, I first turn to a short discussion of word order and case morphology in Basque, Polish and Spanish, as it relates to Experiment 2. The main hypotheses that aim at accounting for cross-linguistic variation in the preference for short dependencies are then presented in Sections 1.2 and 1.3, before reporting and discussing the results in Sections 2 and 3 respectively.

1.1 Grammatical sketches of Basque, Spanish and Polish

Here I focus on three languages, namely Spanish, Polish and Basque, respectively a Romance, a West Slavic and a non-Indoeuropean language isolate. In this study, ditransitive and transitive sentences were presented to the participants and the length of the phrases was manipulated. Sentences were translated and adapted from materials in Section 2.3 in Chapter 3 (see Appendixes 5 to 8).

1.1.1 Case marking

Regarding case marking, Spanish presents a rather poor case marking system. Although it has differential object marking in transitive sentences and the remnants of a former case marking system are still present in the pronominal system –e.g., *él* (Nom.), *le/lo* (Acc.) and *le* (Dat.)– , no case morphology unequivocally distinguishes in Spanish between subjects, objects and indirect objects when NPs are strong pronouns or non-pronominal NPs (Torrego, 1998). So, for example, the indirect object in the set of sentences in 13a and 13b below (*a la señora*, *to the lady*) is marked exactly as the direct object in 12a and 12b (*a la señora*, *the lady*). Overlapping of case marking in the verbal arguments is also dependent on relative animacy, telicity, topicality or dialect, among other factors (Aissen, 2003; Tippets, 2011; Torrego, 1998; Weissenrieder, 1990; Zubizarreta, 1994, *inter alia*). Basque and Polish, in turn, present rich case marking systems. In Polish, the main determinant of case ending is gender, which can be consistently predicted from the phonological form of the nominative case, in which feminine names usually end in *–a*, masculines in a consonant and neuters in *–o, -e or –ę*. When a particular case has more than one possible ending, the appropriate form is consistently predicted by either phonological or semantic factors (Westfal, 1956 and Kottum, 1981, as cited in Siewierska, 1993). The Polish cases relevant for this study are summarized in Table 3.1. The cases used in Standard Basque that are relevant for this study are summarized in Table 3.2 (Trask, 1997).

Case	**Feminine**	**Masculine**	**Neuter**
Nominative	-a (-∅, -i)	-∅ (-a, -o)	-o, -e, -ę
Accusative	-ę (-∅)	- , -a (-ę, -o)	-o, -e, -ę
Dative	-‘e, -i/-y	-owi (-u, -‘e, -i/-y)	-u

TABLE 3.1: **Case marking system (singular endings) in Polish**. Endings between parenthesis are restricted to a very limited set of exceptions. *y/-i* are phonotactically restricted variants of the same ending. The symbol ‘ indicates palatalization of the preceding consonant

Case	Singular	Plural
Ergative	-ak	-ek
Absolutive	-a	-ak
Dative	-ari	-ei

TABLE 3.2: **Case marking system (singular and plural endings) in Basque**. I include the plural endings in this case because ergative plural NPs were used in the transitive sentences in the experiment in Basque

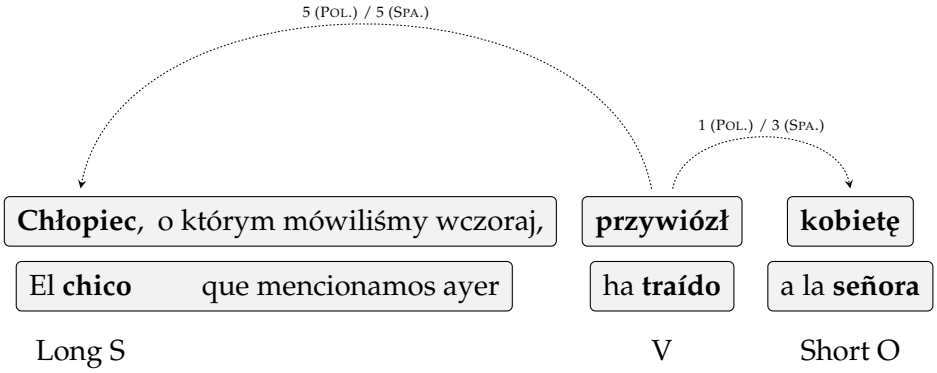
1.1.2 Word order

Basque, Polish and Spanish all allow different word orders in both transitive and ditransitive sentences, so that speakers can choose between propositionally equivalent variants. In the following examples (set of sentences from 12a to 15b), sentences are represented as dependency graphs, to visually illustrate differences in dependency lengths across possible variants. Arches represent dependencies between the verbs and their arguments, measured, as conventionally in corpus studies, in number of words[2]. Shorter dependency lengths obtain when phrasal heads (in bold) are placed adjacent to each other, yielding a *short-before-long* ordering in head-initial Polish and Spanish and a *long-before-short* one in head-final Basque (see Chapter 2). In transitive sentences, S-V-O is the most frequent and basic order in both Polish and Spanish (see 12a set of sentences). In these two languages, when S is long, placing it later in the sentence yields the shortest dependency distance (see 12b set of sentences). In ditransitive sentences, both Spanish and Polish present a S-V-O-IO canonical orders, (see 13a set of sentences). When O is long, the reversed IO-O order yields shorter dependency lengths than its canonical counterpart (see 13b set of sentences). In turn, in Basque, where the basic word orders are S-O-V and S-IO-O-V in transitive and ditransitive sentences respectively, it is the O-S and O-IO orders that yield the shortest dependency length when the O is long (see 14a versus 14b and 15a versus 15a).

(12) **Possible word orders in a transitive sentence with NPs of different length in Polish and Spanish**

a. *Canonical order: long-before-short*

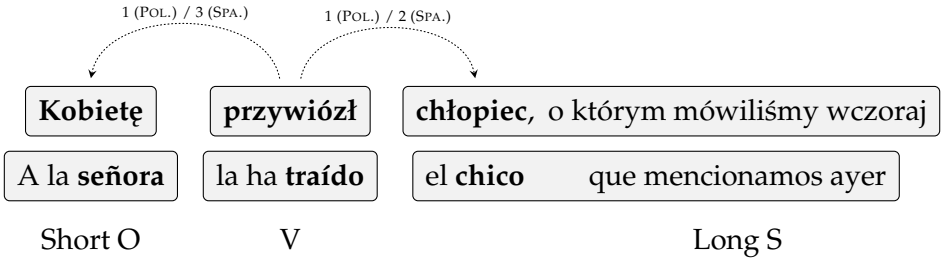
[2]I do not adopt Hawkins' metrics here because of the greater visual simplicity of the representations as dependency graphs. Also, dependency length is calculated by counting all the words between lexical heads, in bold in the examples. For the aims of this study, since word order preferences are a function of relative differences between different licensed linearizations of the same structure, my predictions hold the same whether nouns or determiners are the assumed to be the heads of the relevant phrases (Abney, 1987).



Chłopiec,	o którym	mówiliśmy	wczoraj,	przywiózł	kobietę
El chico	que	mencionamos	ayer	ha traído	a la señora
The boy	that	we mentioned	yesterday	brought	the lady

The boy has brought the lady that (we) mentioned yesterday

b. *Non-canonical order: short-before-long*

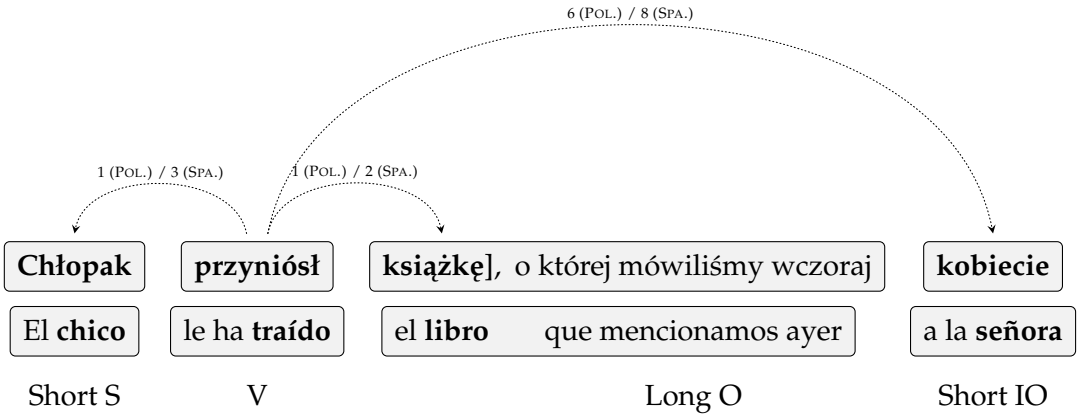


Kobietę	przywiózł	chłopiec,	o którym	mówiliśmy	wczoraj
A la señora	la ha traído	el chico	que	mencionamos	ayer
The lady	brought	the boy	that	we mentioned	yesterday

The boy that we mentioned yesterday has brought the lady

(13) **Possible word orders in a ditransitive sentence with NPs of different length in Polish and Spanish**

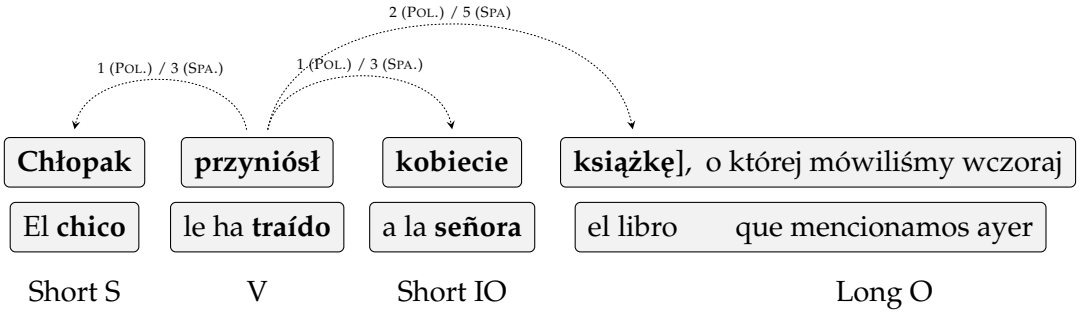
a. *Canonical order: long-before-short*



Chłopak	przyniósł	książkę,	o której	mówiliśmy	wczoraj	kobiecie
El chico	le ha traído	el libro	que	mencionamos	ayer	a la señora
The boy	brought	the book	that	we mentioned	yesterday	to the lady

The boy has brought the lady the book that (we) mentioned yesterday

b. *Non-canonical order: short-before-long*

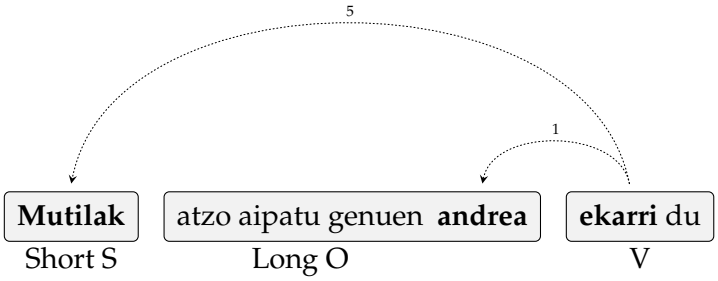


Chłopak	przyniósł	kobiecie	książkę,	o której	mówiliśmy	wczoraj
El chico	le ha traído	a la señora	el libro	que	mencionamos	ayer
The boy	brought	to the lady	the book	that	we mentioned	yesterday

The boy has brought the lady the book that (we) mentioned yesterday

(14) **Possible word orders in a transitive sentence with NPs of different length in Basque**

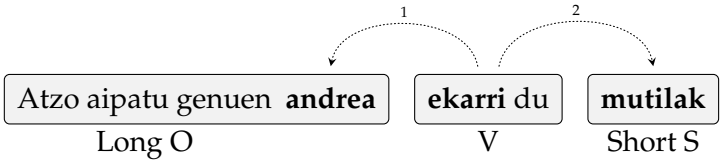
a. *Canonical order: short-before-long*



Mutilak	atzo aipatu genuen andrea	ekarri du
Boy-the	yesterday we mentioned-that lady-the	brought

The boy has brought the lady that (we) mentioned yesterday

b. *Non-canonical order: long-before-short*

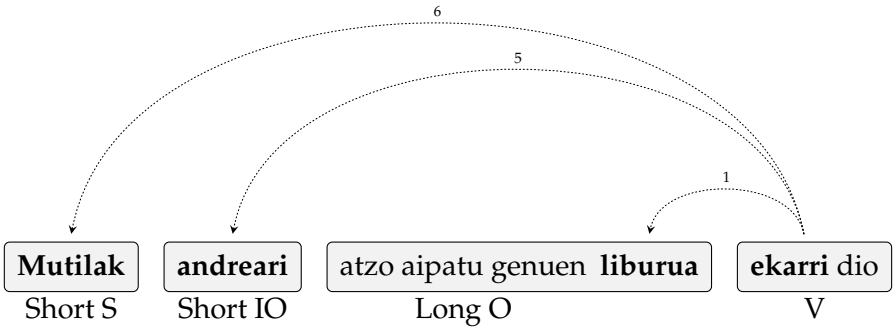


Atzo aipatu genuen andrea		ekarri du	mutilak
Yesterday we mentioned-that lady-the		brought	boy-the

The boy has brought the lady that (we) mentioned yesterday

(15) **Possible word orders in a ditransitive sentence with NPs of different length in Basque**

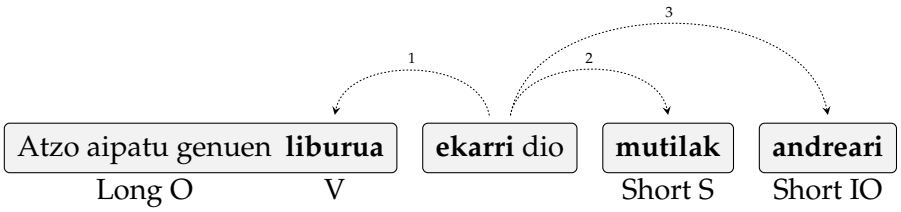
a. *Canonical order: short-before-long*



Mutilak	andreari	atzo aipatu genuen liburua	ekarri dio
Boy-the	lady-to	yesterday we mentioned-that book-the	brought

The boy has brought the lady the book that (we) mentioned yesterday

b. *Non-canonical order: long-before-short*



Atzo aipatu genuen liburua	ekarri dio	mutilak	andreari
Yesterday we mentioned-that book-the	brought	boy-the	lady-to

The boy has brought the lady the book that (we) mentioned yesterday

1.2 DLM and the role of case marking

According to information-theoretic approaches to production (Gibson et al., 2013; Jaeger, 2010), linguistic preferences result from a trade-off between the need to increase the probability of successful recovery of an intended message and the tendency to decrease the effort to produce it (Horn, 1984; Zipf, 1949). Striking a balance across these two principles predicts a preference for reducing redundancy in the cues employed to signal syntactic structure (Fedzechkina et al., 2012; Kurumada and Jaeger, 2015). Experimental evidence for the joint role of robust information transfer and production demands in shaping language comes mainly from phonetic and phonological form (Aylett and Turk, 2004; Bell et al., 2003; Jaeger and Buz, 2017) and lexical items (Mahowald, Fedorenko, Piantadosi, and Gibson, 2013; Piantadosi et al., 2011) and, to a lesser extent, syntax and other linguistic levels (Fedzechkina et al., 2012; Genzel and Charniak, 2003; Gildea and Jaeger, 2015). For example, Jaeger (2010) demonstrated that speakers have a higher preference for mentioning the full form of the

relative pronoun *that* the more information it carries. Also, in an artificial miniature language experiment, Fedzechkina et al. (2013) found that speakers deviated from the received input and reduced the proportion of case-marked nouns more in a fixed word order language, where morphological information was redundant with word order, than in one with relatively flexible word order. Also, it has been shown that case marking tends to be more widespread cross-linguistically for phrases which are less prototypical or predictable (Aissen, 2003; Comrie, 1989; Croft, 1990; Kurumada and Jaeger, 2015).

In this study, I investigated whether rich case morphology reduces the strength of the bias for dependency length minimization. To do so, the strength of the preference for short dependencies in Basque and Polish, on one hand, and in Spanish, on the other, were compared. Minimized dependency length is obtained through word order shifts that lead to phrasal head adjacency (see set of sentences 12a to 15b), which is claimed to make syntactic recognition more reliable and to ensure robust information transmission, by decreasing uncertainty about syntactic and semantic dependencies (Hawkins, 2014; MacWhinney and Bates, 1989). Producing shifted word orders, however, can be resource demanding. Recent studies confirm that, when uttering a sentence, alternative structures compete for selection, restricting the availability of each choice and increasing production effort (Dell and O'Seaghdha, 1994; Hwang and Kaiser, 2014; Myachykov, Scheepers, Garrod, Thompson, and Fedorova, 2013)[3]. Also, a number of studies suggest that competition in production stems from similarity-based interference, as it does in comprehension (Gennari et al., 2012; Humphreys et al., 2016; Slevc, 2011). Furthermore, the effort associated with producing a structure is correlated with the number of possible alternatives present in the language that can express the same proposition (Myachykov et al., 2013).

In languages with rich case morphology such as Basque or Polish, case marking provides clear cues as to which words need to be retrieved or integrated at a given point in a sentence by clearly mapping form and syntactic and semantic function (MacWhinney, Bates, and Kliegl, 1984; Sasaki and MacWhinney, 2006), which would mitigate the processing cost of long dependencies. Resorting to cognitively costly shifted orders would be deemed inefficient, since it increases production effort unnecessarily (Fedzechkina et al., 2013). In contrast, in languages with poor or lacking a case marking system such as English or Spanish, speakers will necessarily rely more strongly on word order to signal syntactic and semantic dependencies, regardless of production cost, in so far as no other strategy is available (Kempe and MacWhinney, 1999; MacWhinney and Bates, 1989). Languages with case morphology such as Basque or Polish will thus show a greater tolerance for long dependencies than

[3]V. S. Ferreira, 1996 argues that the presence of syntactic choice eases production demands: English speakers produce sentences faster and make fewer mistakes when prompted with verbs that allow alternative structures (e.g. *give*) than with verbs that did not allow them (e.g. *donate*). However, as Myachykov et al. (2013) point out, the sets of verbs used in Ferreira's study differed in terms other than just syntactic flexibility, such as frequency and number of syllables: *give* is shorter and more frequent than *donate*. Also, the participants were under time pressure, and time constraints are related to more incremental language planning scope, which can impact the availability of competing structures (F. Ferreira and Swets, 2002)

languages with poor or absent case morphology such as English or Spanish, in any type of dependency and without any disadvantages in terms of processing efficiency.

1.3 DLM and the role of word order based-expectations

Alternatively, cross-linguistic differences in the strength of the bias for short dependencies could emerge due to distributional properties that stem from specific grammatical properties (in this case, linear word order) that allow speakers and listeners to build expectations about upcoming linguistic elements. The cognitive cost of retrieving a given linguistic item decreases as its conditional probability increases (Hale, 2001; Levy, 2008). In formal terms, processing difficulty for a given linguistic element (e. g. a word or a part of speech) is based on its Shannon information value in context, that is on the log inverse probability of the word in a given context (Hale, 2001; Levy, 2008; N. J. Smith and Levy, 2013). This means that processing is facilitated if a given linguistic form is more predictable from the information given by another one previously encountered, independently (or, according to certain accounts such as Levy (2006) and Levy et al. (2008) because) of the distance between them, contrary to predictions from memory-driven locality effects. Indeed, in comprehension, faster reading times have been found for sentences with a higher number of dependents before the verb in German (Konieczny and Döring, 2003) and Hindi (Husain et al., 2014; Vasishth and Lewis, 2006). In Japanese increasing the distance across dependencies does not correlate with slower reading times (Nakatani and Gibson, 2010). It has been suggested that these so-called *antilocality* effects are more common in head-final languages such as Hindi or Japanese, whereas *locality* effects predominate in head-initial languages such us Chinese, English or Russian (Levy et al., 2013). However, surprisal effects have also been found in comprehension studies in English (Levy et al., 2013; Staub, 2010) or Chinese (Jäger, Chen, Li, Lin, and Vasishth, 2015).

I investigated whether expectation-based effects can also play a role in production data. I was interested in whether, as in comprehension, the preference for short dependencies in production varies depending on the level of predictability of the upcoming element. Furthermore, I was interested in the effect of predictability regarding the degree of expectation of a given part of speech across languages. Word order-based expectations will have different levels of certainty not only across constructions but also across languages depending on the distributional regularities stemming from the licensed grammatical orders in those languages. This will lead to cross-linguistic divergences in the profiles of processing difficulty of the same constructions. My working hypothesis is that, also in production and across languages, the more expected a given part of speech is, the smaller the cognitive load of processing the long dependency that it resolves. To investigate this, first three completion studies were conducted (in Basque, Polish and Spanish) in order to derive approximate estimations for the parts of speech that I was interested in measuring, that is for those parts of speech that in the materials (transitive and ditransitive sentences) appear after a long dependency (see Section 2.3 below).

Sentence completion studies have been previously used to derive predictions of upcoming elements in Levy et al. (2013), Husain et al. (2014), Jäger et al. (2015) and Safavi et al. (2016).

1.3.1 Sentence completion studies in Basque, Polish and Spanish

In three brief web-based pre-tests conducted at the Ibex Farm platform, the probability of a given syntactic category was calculated in relation to its preceding context (Kutas and Hillyard, 1984) in Basque, Polish and Spanish. Two sets of transitive and ditransitive sentences (12 sentences, total) were randomly selected from the items used in the main experiment (see Section 2.3). The relevant syntactic category whose probability value I was interested in estimating (O, IO and V) was erased (see Appendices 2, 5 and **??**). Participants (n = 46, 50 and 50 in Basque, Polish and Spanish respectively) had to fill in the blank in the incomplete sentence with any word(s) they judged to be most appropriate. Cloze probability of S, IO or O given the previous context was measured by calculating the proportion of times that a given element (O, IO and V) was produced over all responses (Tayor, 2015). Crucially, I sought to obtain an estimate of the probability of those syntactic categories that followed a long dependency in the critical conditions in the materials employed in the production experiments (see Section 2.3), that is after S- in Basque (Long-O condition) and in Polish and Spanish (Long-S condition) in transitive sentences and after S-IO- in Basque (Long-O condition) and S-V-O- in Polish and Spanish (Long-O condition) in ditransitive sentences[4]. In the sentences used in the completion tasks no relative clauses were employed. As stated above, I was interested in determining whether differences in the strength of the expectations for upcoming constituents would modulate the preference for short dependencies across languages. To do so, I first needed to estimate the differences in expectations across Basque, Spanish and Polish, the three languages tested in this study. See Table 3.3 for an example of a set of the type of sentences. Results are reported below. The resulting data were analyzed using logistic regressions with a binomial link function (see Tables 3.4 and 3.5).

In the sentences employed in the cloze task, dative clitics were included in the Spanish materials, in order to mimic real use. This contrasts with materials in the cued-recall production experiment itself, where the verb was provided in its infinitive form (see Section 2.7.2). Clitic doubling in Spanish is described as optional whenever the NP in IO position bears the thematic role of goal and is placed in a non-topical position, and compulsory when it bears the thematic role of beneficiary, location or possessor (Bickford, 1985; Demonte, 1995; Jaeggli, 1982). Approximately half of the experimental items used in the experiment belong to these three last categories (17 out of 30). Also, although Spanish is described as a language

[4]The RC precedes the head in Basque and follows it in Spanish and Polish. This means that the long dependency appears before the O in both transitive and ditransitive sentences in Basque. In Spanish and Polish, in ditransitive sentences the relative clause follows the O. In these languages, in transitive sentences, in turn, a long dependency occurs when the S, and not the O, is followed by a long RC

Set of sentences	**Examples**
Transitive sentences	(a) S-...: Modeloek _____. (b) S-...: La modelo _____. (c) S-...: Modelka _____. *The models* _____. (Bas que)/ *The model* _____. (Polish/Spanish)
Ditranstive sentences	(a) S-IO...: Ainhoak lankideari _____. (b) S-V-O...: Clara le ha explicado el problema _____. (c) S-V-O...: Marta wyjaśniła problem _____. *Ainhoa to the colleague* _____. (Basque) / *Clara/Marta CL/∅ explained the problem* _____. (Polish/Spanish)

TABLE 3.3: **Example set of experimental transitive and ditranstive sentences in the sentence completion tasks in (a) Basque, (b) Polish and (c) Spanish in Experiment 2**

with optional clitic doubling, several corpus studies show that in spoken contemporary peninsular Spanish, clitic doubling of IO is the predominant realization (Flores, 2004; Silva-Corvalán, 1984), reaching up from 70% to 90% of all corpus occurrences of indirect objects. In Basque Spanish used by the participants in this study (the diatopical variety of Spanish spoken by Basque bilinguals and monolinguals who have lived most of their lives in the Basque Country), clitic doubling is quite a widespread phenomenon, since it is possible with all animate objects, both IO and O, unlike in some other varieties of Spanish where the doubling of O by clitics is quite restricted (Landa, 1995). Due to all this, although in the cued-production task employed in Experiment 2 dative clitics were not always produced, probably due to its absence in the verb present as a cue to the participants, they were included in the cloze-task materials.

In Spanish and Polish transitive sentences, the probability of a V appearing after a S was of .89 [CI = .93 – .82] and .87 [.91 – .79] respectively, so that the presence of an utterance-initial animate NP greatly narrowed down the range of possible continuations. In contrast, in Basque, the type of constituents that could follow the S was less restricted, being an O in only .60 of the cases [CI: .68 – .51]. As shown in Table 3.4, in transitive sentences, the uncertainty about the upcoming constituent was found to be greater in Basque than in either Spanish or Polish. No significant difference was found between Spanish and Polish ($p < .9$). In consequence, if the degree of predictability of a given syntactic category mitigates the processing cost induced by long dependencies, the tendency to shorten dependency length should be stronger in Basque than in Polish or Spanish. Polish and Spanish should not differ from each other. The cross-linguistic difference in the cloze value of the syntactic category following the S is not due to the ambiguity of the case ending *-ak* in Basque (see Table 3.2), which can be interpreted as an aboslutive-marked plural NP, that is as S of an upcoming intransitive clause (Yetano, Duñabeitia, and Laka, 2012). This possibility was discarded by using unambiguous plural ergative-marked NPs (*-ek*) both in the completion study and in the cued recall production task (see Section 2.3).

In ditransitive sentences in Basque, for which S-IO-O-V is argued to be the canonical order (Elordieta, 2001), I confirmed a strong probability for an upcoming O phrase and thus a

ditransitive lexical frame when an ergative-marked noun phrase is encountered at the beginning of a sentence followed by a dative-marked one, that is after a S-IO-... sequence (.71, CI = .78 – .62). In turn, in both Polish and Spanish, where the most frequent word order is S-V-O-IO, the probability of an IO after encountering a S and the corresponding V and O was weaker than in Basque (.58 [.49 – .67] and .58 [.50 – .66] in Polish and Spanish respectively). The expectation of the syntactic categories that in sentences similar to the ditransitive materials follow a long dependency was significantly greater in Basque than in Polish and Spanish and it did not differ across these two ($p > .9$), as shown in Table 3.5. If expectation-based accounts of processing cost also apply in production, then in ditransitive sentences, as in transitive sentences, Basque should differ from both Polish and Spanish. Now, Spanish and Polish should show a stronger preference for short dependencies than Basque, since the expectation of the relevant upcoming syntactic category is stronger in Basque than in Polish or Spanish. Again, Polish and Spanish should not differ from each other.

	Estimate	SE	z-value	p-value
Intercept	.399	.174	2.292	02 *
Spanish	1.480	.297	4.989	< .001 ***
Polish	1.460	.302	4.835	< .001 ***

TABLE 3.4: **Model results from the logistic regression for the sentence completion study in Basque, Polish and Spanish (transitive sentences) in Experiment 2**

	Estimate	SE	z-value	p-value
Intercept	.876	.188	4.652	< .001 ***
Spanish	-.536	.251	-2.138	.03 *
Polish	-.552	.258	-2.137	.03 *

TABLE 3.5: **Model results from the logistic regression for the sentence completion study in Basque, Polish and Spanish (ditransitive sentences) in Experiment 2**

2 Experiment 2

2.1 Overview of the experiment

In this study, I investigated the likelihood that speakers produce more shifted word orders when a constituent is long than when short, in both transitive and ditransitive sentences in Spanish, Polish and Basque. Specifically, I sought to test whether the preference for short

dependencies that derives from the use of shifted orders when one constituent is long varies across languages as a function of either case-marking or of word order-based expectations. To do so, a cued-recall production task was conducted, where participants were presented with unordered phrases on a computer screen and were asked to construct sentences using these phrases. Similar sentences were employed in all three languages, and the length of the phrases and thus of the resulting dependencies (short vs. long) was manipulated by adding relative clauses to different constituents.

Let me briefly review the predictions made by the hypotheses put forth in Sections 1.2 and 1.3 above. One, if case marking reduces the processing cost of long dependencies by reducing uncertainty about the intended parsing, then Basque and Polish, where grammatical function assignment is achieved through consistent case marking, will show an overall weaker pressure for producing short dependencies in all type of sentences compared to Spanish, a language with scant case morphology. Second, it is possible, however, that cross-linguistic differences emerge from prior experience with language-specific word orders, which in turn inform expectations about upcoming linguistic material. If, in production as in comprehension, the cognitive cost of a certain item is modulated by its level of predictability, word orders that shorten dependency length will not be advantageous in conditions where certain grammatical properties of the language(s) at stake generate strong expectations about incoming material. If this is so, based on the cloze values reported in Section 1.3.1, in transitive sentences, one would expect a stronger preference for short dependencies in Basque than in Polish or Spanish. In contrast, in ditransitive sentences, the inverse pattern should be expected, that is, a weaker preference for short dependencies in Basque than in Polish or Spanish. Polish and Spanish should not differ from each other in any case.

2.2 Participants

42 students participated in this experiment: 18 native speakers of Spanish (16 females, mean age = 25.56), all of them undergraduates at the University of the Basque Country, and 24 native speakers of Polish (21 females, mean age = 22.04), students at the Adam Mickiewicz University. I compared their data to that obtained in Experiment 1, where 25 native speakers of Basque (n = 1 excluded due to technical error, 15 females, mean age = 19.39) participated (see Section 2.2). Participants in Experiment 1 were L1Basque-L2Spanish bilinguals, due to the sociolinguistic situation of the Basque Country, where virtually no monolingual speakers are to be found (Department of Education and Culture, 2013a). High proficient L1Spanish-L2Basque and L1Polish-L2Spanish bilingual speakers were therefore recruited (see Table 3.6), in order to reduce the possibility that any differences in the strength of the bias for DLM across languages could be attributed to the effect of bilingualism on working memory or attention (Calvo, Ibañez, and García, 2016). The experiment was conducted in the participants' respective L1s, Spanish and Polish. Although they were all highly proficient in their L2, there were inevitable differences with regard to the AoA of their L2 (2.5 [SD = 1.4]

and 15.6 [SD = 3.9] for Spanish and Polish participants respectively) (see Appendices 1 and B for the linguistic questionnaire completed by Spanish and Polish participants, respectively). Participants all signed an ethical consent form and they were compensated economically for their participation.

	Basque (L1 = Spanish)	**Spanish (L1 = Polish)**
Listening	6.72 (.46)	5.79 (1.14)
Speaking	6.22 (.73)	5 (1.22)
Reading	6.61 (.50)	5.58 (1.10)
Writing	6.28 (.67)	4.96 (1.08)
Mean	6.46 (.52)	5.33 (1.04)

TABLE 3.6: **Self-rated proficiency level scores (SD in brackets) for participants' L2s, Basque and Spanish, In Experiment 2**. *Self-assessed language proficiency is rated on a 7-point scale: 7 = Native-like level; 6 = High level; 5 = Medium-high level; 4 = Medium level; 3 = Medium-low level; 2 = Low level; 1 = No knowledge*

2.3 Materials

24 transitive and 30 ditransitive experimental sentences (translated and adapted to Spanish and Polish from the materials in Basque used in Experiment 1) (see Section 2.3) and 56 fillers were created for each language (see Appendices 5 to 10). Three different versions of each sentence were created by attaching a relative clause to (1) the subject (for transitives, Long-S) and the indirect object (for ditransitives, Long-IO); (2) the object (Long-O); or (3) neither of them (All-Short) (see Table 3.7 for transitive sentences and Table 3.8 for ditransitive sentences). This manipulation yielded the three main Length conditions of the experiment. In the transitive sentences, the relative clause had a mean length of 11.9 syllables (SD = 2.5) in Basque, 12.2 (SD = 2.5) in Polish and 14.5 (SD = 3.2) in Spanish. In the ditransitive sentences, it had a mean length of 10.5 (SD = 2) syllables in Basque, 10.5 (SD = 2.3) in Polish and 12.2 (SD = 2.5) in Spanish. In order to avoid effects due to reading habits, the position of the constituents on the screen was counterbalanced: S-O vs. O-S in the transitive sentences and IO-O vs. O-IO in the ditransitive sentences, where the first phrase is located to the left of the second, in a left-to-right reading order.

There were some minor differences across materials (see Tables 3.7 and 3.8) to accommodate differences in vocabulary, lexical biases, grammatical features, and cultural background between Basque, Spanish and Polish. First, in Spanish transitive sentences, an accusative clitic needs to be added when the object is fronted (see examples 13a and 13b in Section 1.1) and, in ditransitives, there is optional dative clitic doubling, that is, an unstressed dative pronoun can co-occur with a co-referential IO (shown in brackets in examples (c) in Table 3.8). In Polish and Basque, in turn, the inflected verb carries only agreement morphemes. In order to avoid any possible effect of the presence or absence of the clitic in

Spanish, the main verb was always presented in infinitive in this language, unlike in Polish and Basque where the inflected form was employed. Second, in Polish, unlike in Spanish or Basque, two commas or a comma and a dot obligatorily demarcate the beginning and end of a relative clause. The experimtental materials only contained the first comma, between the antecedent and the relative pronoun, without a second comma or a dot at the end of the relative clause, in order to avoid that the results obtained could be biased by orthographical conventions. Third, in Polish, unlike in Spanish and Basque, the relative pronoun agrees with its antecedent in gender and number and bears the case marking required by the verb in the relative clause, and thus different forms of the same relative pronoun were used when required (see Table 3.8 for an example, where the forms *która* (femenine) and *który* (masculine) are used). Fourth, in Basque transitive sentences the head of the subject phrases appeared in plural (see Table 3.7). This was due to the fact the ergative singular forms are homophonous with absolutive plural nouns, and I wanted to reduce the possibility that any type of ambiguity during functional encoding could impact structural preferences (see Table 3.7, where *andreek*, "ladies", is used) [5]. Lastly, in Polish some verbs had to differ from their Spanish and Basque counterparts because they required their complements to have a case other than accusative (for an example see Table 3.7).

Phrase length condition	**Example**
All-Short	(a) Langileek politikaria aipatu dute
	(b) Robotnik wspomniał związkowca.
	(c) El trabajador ha mencionado al político.
	The worker(s) mentioned the politician.
Long-S	(a) [Asanbladan denon aurrean hitz egin duten] langileek politikaria aipatu dute.
	(b) Robotnik, [który wygłosił przemowę na zebraniu], wspomniał związkowca.
	(c) El trabajador [que dio un discurso en la asamblea] ha mencionado al político.
	The workers [that spoke (in front of everyone) at the meeting] mentioned the politician.
Long-O	(a) Langileek [asanbladan denon aurrean hitz egin duten] politikaria aipatu dute.
	(b) Robotnik wspomniał związkowca, [który wygłosił przemowę na zebraniu].
	(c) El trabajador ha mencionado al político [que dio un discurso en la asamblea].
	The workers mentioned the politician [that spoke (in front of everyone) at the meeting]

TABLE 3.7: **Example set of experimental transitive sentences in (a) Basque, (b) Polish and (c) Spanish in Experiment 2**

[5]Subjects NPs in the Basque transitive sentences in the experimental stimuli carry plural number, which has been shown to affect word order preferences in a number of languages (Aissen, 2003). However, if this had any effect at all, it would have biased against and not in favor of my predicted outcome for Basque - that is, a *long-before-short* preference. This is so because plural phrases tend to precede singular ones, as shown for English in a corpus study by Bresnan et al., 2007 (p. 17), where plural recipients trigger the double object construction (e. g. *gave the children [plural recipient] a toy*), whereas plural themes trigger a PP dative (e. g. *gave the toys to the chid*). This means that the use of plural subjects in the materials, if at all, would have reinforced the tendency of towards the canonical S-O-V orders in transitive sentences.

Phrase length condition	Example
All-Short	(a) Argitaratzaileak itzultzaileari liburua eskaini dio. (b) Redaktor ofiarował książkę tłumaczowi. (c) El editor (le) ha ofrecido el libro al traductor. *The editor offered the book to the translator.*
Long-IO	(a) Argitaratzaileak [iaz arrakasta handia zuen] itzultzaileari liburua eskaini dio. (b) Redaktor ofiarował książkę tłumaczowi, [który odniósł w ubiegłym roku olbrzymi sukces]. (c) El editor (le) ha ofrecido el libro al traductor [que tuvo tanto éxito el año pasado]. *The editor offered the book to the translator [that was a great success last year].*
Long-O	(a) Argitaratzaileak itzultzaileari [iaz arrakasta handia zuen] liburua eskaini dio. (b) Redaktor ofiarował książkę, [która odniósła w ubiegłym roku olbrzymi sukces] tłumaczowi. (c) El editor (le) ha ofrecido el libro [que tuvo tanto éxito el año pasado] al traductor. *The editor offered the book [that was a great success last year] to the translator.*

TABLE 3.8: **Example set of experimental ditransitive sentences in (a) Basque, (b) Polish and (c) Spanish in Experiment 2**

2.4 Procedure

Experiment 2 was based on a previous design employed by Stallings et al. (1998) and Yamashita et al. (2001). The procedure was identical to the one used in Experiment 1 (see Section 2.4). Six pseudo-randomly ordered lists were constructed. Experimental sentences of the same type (e. g., transitives or ditransitives) were never presented next to each other. The experiment was presented under the guise of a memory task: 11 fillers were repeated, and participants responded whether they had previously appeared through the experiment. In each trial, participants pressed the spacebar, triggering the display of four boxes containing sentence constituents. Participants prepared a sentence and pressed the spacebar when ready. After 500 milliseconds, an arithmetic operation with two possible solutions was presented with the aim of preventing covert rehearsal (Yamashita and Chang, 2001). Participants solved the problem and pressed the space bar. Then, the verb of the sentence to be recalled was presented as a cue, and participants produced the sentence.

2.5 Scoring

Responses were scored in terms of word order as shifted versus non-shifted. Sentences were scored as *non-shifted* whenever their constituents were recalled in their canonical order (transitive sentences: S-O in all three languages; ditransitive sentences: O-IO in Polish and Spanish and IO-O in Basque). They were scored as *shifted* when their constituents were recalled in a non-canonical order (transitive sentences: O-S in all three languages; ditransitive sentences: IO-O in Polish and Spanish and O-IO in Basque). I scored as miscellaneous sentences partial or no recalls, ungrammatical sentences, or when participants swapped the original length, function or number of any of the constituents, used an indefinite NP or employed non-target structures. In the transitive sentences, a total of 13, 20 and 19 data points were removed from the analysis in Basque, Polish and Spanish respectively. Similarly, in ditransitive sentences, a total of 14, 31 and 50 data points were thus discarded from further

analysis in Basque, Polish and Spanish respectively. Changes in tense were accepted as well as synonyms and minor deletions, changes and additions (in transitive sentences: n = 53 in Basque; n = 95 in Polish; n = 98 in Spanish; in ditransitive sentences: n = 72 in Basque; n = 114 in Polish; n = 116 in Spanish). In Polish sentences, 7 data points in the ditransitive sentences and 1 data point in the transitive sentences consisted in relative clause extrapositions, thas is the relative clause modifying the NP was produced postverbally, separate from its antecedent. Extraposed relative clauses are found cross-linguistically (Konieczny, 2000 and Uszkoreit et al., 1998 in German; Francis, 2006 in English), and they yield shorter dependencies than its canonical counterpart (Arnold et al., 2000; Wasow, 2002, *inter alia*). I thus included them in my analysis as instances of shifted orders. In ditransitive sentences, scoring resulted in 703 (98%) target responses in Basque, 683 (95.7%) in Polish and 490 (90.7%) in Spanish. In transitive sentences, scoring resulted in 563 (97.7%) target responses in Basque, 555 (96.5%) in Polish and 413 (95.6%) in Spanish.

2.6 Analysis

The maximum random effect structure justified by the data was used, using model comparison between all the models that converged and did not contain correlations between the random effects equal to 1/-1 (Baayen et al., 2008). I report p-values estimated by the R-package lmerTest (Kuznetsova, Brockhoff, and B.Christensen, 2017). All analyses were conducted using R (R Core Team, 2014) and the lme4 package (D. M. Bates et al., 2014). No analysis of the effect of constituent length in Spanish and Polish transitive sentences nor any statistical comparison across languages was conducted due to the limited and zero-number observations per cell (Jaeger, 2011).

First, I analyzed the effect of dependency length on Spanish and Polish separately. The results for Basque are reported in Chapter 2. For the ditransitive sentences, I fitted my data to two logit mixed effects models – one for Spanish and the other for Polish –, both with Length and NP Position on the Screen as a predictors, and by-participant and by-item random intercepts and by-participant random slopes. Next, in order to estimate whether there were cross-linguistic differences in the effect of Length, my data was fitted to a logit mixed effects model, with Length, Language (Basque vs. Polish vs. Spanish) and NP Position on the Screen and their interaction as fixed predictors, as well as by-item and by-subject random intercepts. Length was Helmert-coded, and NP Position on the Screen was centered. Specifically, I compared the number of shifted orders produced when the constituents were long to when they were short, regardless of the direction of shift itself (either *short-before-long* or *long-before-short*). Note that when constituents are long they need to be placed in non-canonical shifted positions so that dependency length is minimized (see the set of sentences 12a - 15b presented in Section 1.1).

2.7 Results

2.7.1 Transitive responses in Polish and Spanish

For the first time, I experimentally tested the effect of dependency length in transitive sentences in two head-initial languages that allow scrambling between S and O. In Polish and Spanish, both head-initial languages, when O is long, the maximally efficient *short-before-long* order overlaps with the canonical S-V-O order, leaving little room to explore the effect of dependency length minimization in this case as compared to when O is short. Crucially, a higher number of shifted *short-before-long* non-canonical OVS orders could be produced when S is long than when short, so that dependency length is minimized (see examples 12a in Section 1.1). OVS orders in transitive sentences are allowed in Spanish and Polish and have been reported in previous psycholinguistic research in Spanish (Prat-Sala and Branigan, 2000) and a corpus study in Polish (Siewierska, 1993). However, I failed to find any effect of dependency length in either Polish or Spanish transitive sentences. Spanish speakers did not produce any shifted word orders in the Long-S condition, and Polish speakers produced only 4,6% shifted orders more than in the All-Short condition (see Table 3.9 and Figure 3.1). This runs contrary to previous results from transitive sentences in all head-final languages tested so far with the same procedure as reported here (Dennison, 2008; Ros et al., 2015; Yamashita and Chang, 2001), where orders that minimize dependency length have been consistently reported. Note that in head-final languages the critical condition is Long-O, and not Long-S as in head-initial languages: When S is long, one would not expect to find any effect for the same reason outlined above for Polish and Spanish. In this case, the maximally efficient *long-before-short* order corresponds to the more frequent canonical order SOV (see example 14a in Section 1.1). In Experiment 1, I found that shifted orders were produced more often in Basque when O was long compared to when it was not, in the same proportion as in ditransitive sentences (approx. 19%; see Table 3.9 and Figure 3.1). In turn, Japanese and Korean speakers produced shifted orders 27.5% and 29.4% more often when when the object was long than when it was short.

Language	**Length condition**		
	All-Short	Long-S	Long-O
Basque	7 (3.6%)	10 (5.5%)	37 (19.7%)
Spanish	0 (0%)	0 (0%)	1 (.1%)
Polish	3 (1.6%)	11 (6%)	3 (.2%)

TABLE 3.9: **Raw data and percentages (in brackets) of shifted O-S responses in transitive sentences in Basque, Polish and Polish transitive sentences in Experiment 2**

Language	NP Position on the Screen	Length condition		
		All-Short	Long-S	Long-O
Basque	*Inconsistent with shift direction*	1 (1.1%)	1 (1.1%)	16 (17%)
	Consistent with shift direction	6 (6.3%)	9 (9.9%)	21 (22.3%)
	Miscellaneous	0	9	4
Spanish	*Inconsistent with shift direction*	0 (0%)	0 (0%)	0 (0%)
	Consistent with shift direction	0 (0%)	0 (0%)	1 (1.6%)
	Miscellaneous	3	9	11
Polish	*Inconsistent with shift direction*	0 (0%)	6 (6.6%)	1 (1.1%)
	Consistent with shift direction	3 (3.2%)	5 (5.5%)	2 (2.2%)
	Miscellaneous	3	7	10

TABLE 3.10: **Raw data and percentages (in brackets) of shifted O-S responses in transitive sentences relative to position of O and S on the screen in Basque, Spanish and Polish transitive sentences in Experiment 2**

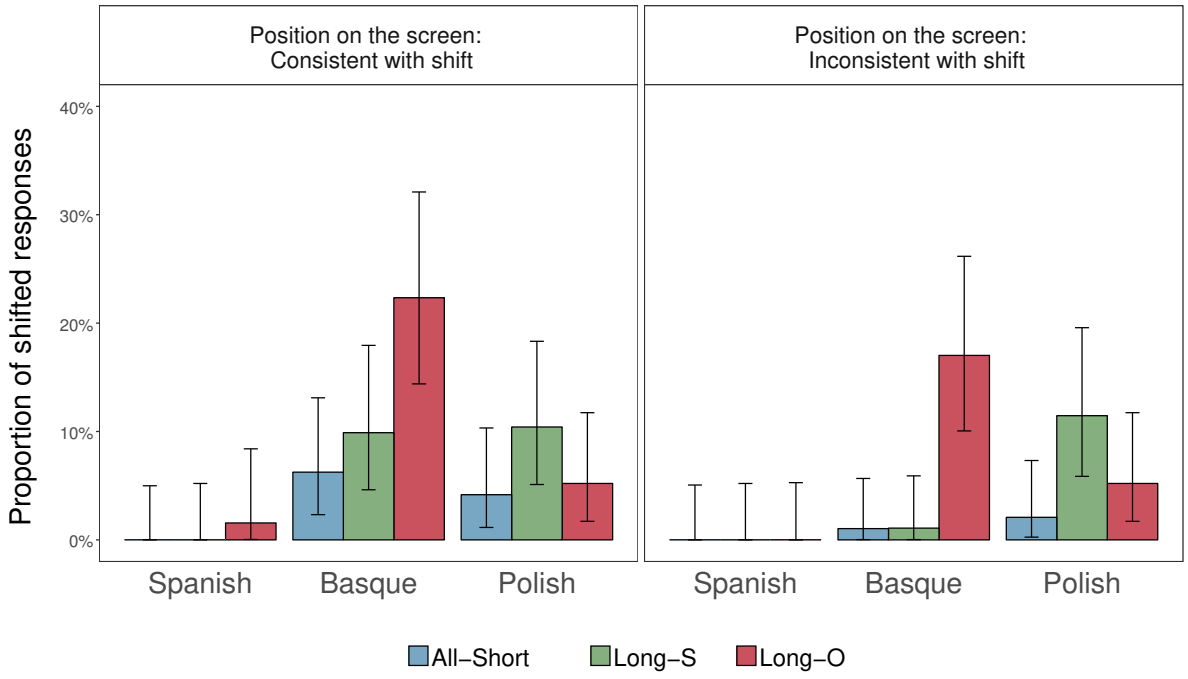


FIGURE 3.1: **Proportion of shifted responses in Basque, Polish and Spanish transitive sentences** (note the different proportion scale than in Figure 3.2). The error bars show 95% confidence intervals. Position on the Screen was included for consistency reasons with respect to Figure 3.2

2.7.2 Ditransitive responses in Polish and Spanish

In Polish ditransitive sentences, there was a preference for S-V-O-IO canonical orders, so that 63.5% of the produced orders were canonical ($\hat{\beta}$ = -.606, z = -2.325, p = .02). Crucially, Polish speakers shifted constituents to later positions more often when they were long than when short, yielding a *short-before-long* preference (Long-O: $\hat{\beta}$ = 1.883, z = 12.422, p < .001; Long-IO: $\hat{\beta}$ = -1.649, z = -6,463, p < .001). However, there was an interaction between NPs Position on the Screen and Long-IO ($\hat{\beta}$ = .739, z = 2.934, p = .002). This was due to the differential effect of NP Position on the Screen on word order preferences. When O appeared previous to the IO on the screen, in a IO-O reading order consistent with the direction of shift in Spanish, the effect of dependency length was stronger when all constituents were short (consistent with shift: 59.5% vs. inconsistent with shift: 13.7%) than when the IO was long (consistent with shift: 1.9% vs. inconsistent with shift: 2.7%). The effect of dependency length when the IO was long is thus difficult to interpret. Nevertheless, the interaction between Position on the Screen and Long-O was non-significant ($\hat{\beta}$ = .119, z = .915, p = .4) (see Figure 3.1 and Table 3.12). Statistical results are reported in Table 3.13.

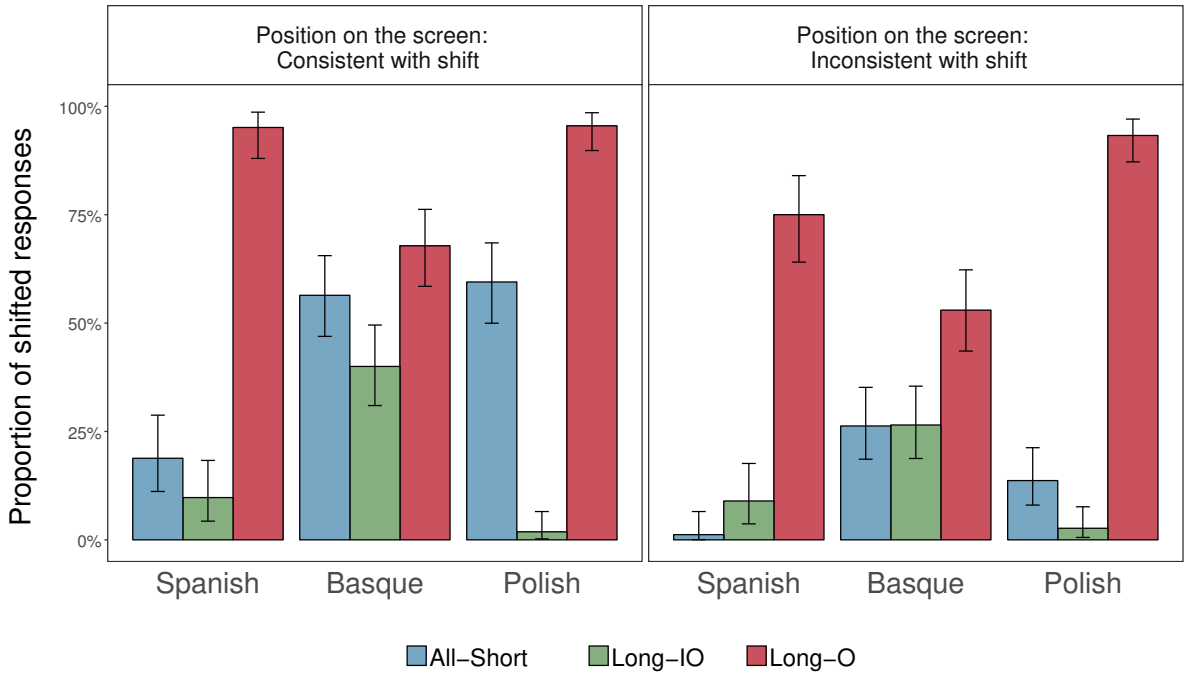


FIGURE 3.2: **Proportion of shifted responses in Basque, Polish and Spanish ditransitive sentences** (note the different proportion scale than in Figure 3.1). The error bars show 95% confidence intervals

Language	**Length condition**		
	All-Short	Long-IO	Long-O
Basque	97 (41.3%)	77 (33.2%)	140 (60.3%)
Spanish	17 (10%)	15 (9.4%)	138 (85.2%)
Polish	81 (36.5%)	5 (2.3%)	217 (96.4%)

TABLE 3.11: **Raw data and percentages (in brackets) of shifted O-IO (Basque) and IO-O (Spanish, Polish) responses in ditransitive sentences in Basque, Spanish and Polish ditransitive sentences in Experiment 2**

Language	**NP Position on the Screen**	**Length condition**		
		All-Short	Long-IO	Long-O
Basque	*Inconsistent with shift direction*	31 (26.3%)	31 (26.5%)	62 (53%)
	Consistent with shift direction	66 (56.4%)	46 (40%)	78 (67.8%)
	Miscellaneous	4	4	6
Spanish	*Inconsistent with shift direction*	1 (1.2%)	7 (9%)	60 (75%)
	Consistent with shift direction	16 (18.8%)	8 (9.8%)	78 (95.1%)
	Miscellaneous	12	20	18
Polish	*Inconsistent with shift direction*	12 (13.7%)	3 (2.7%)	111 (95.7%)
	Consistent with shift direction	69 (59.5%)	2 (1.9%)	106 (97.3%)
	Miscellaneous	5	15	10

TABLE 3.12: **Raw data and percentages (in brackets) of shifted O-IO (Basque) and IO-O (Spanish, Polish) responses relative to position of O and S on the screen in Basque, Spanish and Polish ditransitive sentences in Experiment 2**

	Estimate	SE	z-value	p-value
Ditransitive sentences (N = 683)				
Intercept	-.606	.260	-2.325	.02 *
Length: Long-IO	-1.649	.255	-6.463	< .001 ***
Length: Long-O	1.883	.152	12.422	< .001 ***
NP Position	-.448	.196	-2.287	.02 *
Length (Long-IO) * NP Position	.739	.2519	2.934	0.002 **
Length (Long-O) * NP Position	.119	.131	.915	.360

TABLE 3.13: **Logit mixed model analysis for ditransitive sentences in Polish in Experiment 2**

In Spanish, a strong preference to produce canonical S-V-O-IO orders was observed when all constituents were short ($\hat{\beta}$ = -1.259, z = -3.496, p < .001). Importantly, when O was long,

Spanish speakers shifted it to sentence-final positions, yielding a *short-before-long* IO-O shifted order ($\hat{\beta}$ = 1.913, z = 9.725, p < .001). As in the case of Polish, there was an interaction between NP Position on the Screen and Length ($\hat{\beta}$ = 771, z = 2.528, p = .01), triggered by the fact that, when the O appeared previous to the IO on the screen, in a IO-O reading order consistent with the direction of shift in this language, a higher number of shifted orders was produced in the All-Short condition compared to when the position of the constituents was inconsistent with the direction of shift (a O-IO order) (consistent with shift: 18.8% vs. inconsistent with shift: 1.2%), whereas the number of shifted orders in the Long-IO condition remained relatively stable across conditions (consistent with shift: 9.8% vs. inconsistent with shift: 9%) (see Figure 3.1 Table 3.12). Again, as in Polish, the interaction between Position on the Screen and Long-O was non-significant. Statistical results are reported in Table 3.14.

	Estimate	SE	z-value	p-value
Ditransitive sentences (N = 490)				
Intercept	-1.259	.3601	-3.496	< .001 ***
Length: Long-IO	.330	.304	1.086	.278
Length: Long-O	1.913	.197	9.725	< .001 ***
NP Position	-.938	.232	-4.036	< .001 ***
Length (Long-IO) * NP Position	.771	.305	2.528	.02 *
Length (Long-O) * NP Position	-.111	.146	-.757	.449

TABLE 3.14: **Logit mixed model analysis for ditransitive sentences in Spanish in Experiment 2**

In sum, in ditransitive sentences, all three languages tended to shorten dependency length by shifting constituents and placing phrasal heads adjacent to each other. In Basque, when the O was long, a preference for *long-before-short* shifted and verb-medial orders was found, maximally minimizing dependency length (see Section 2.7 in Chapter 2). For Spanish and Polish, I replicated the *short-before-long* preference found for other sentence types in previous experimental studies in English for NP PP constructions (Francis, 2006; Stallings and MacDonald, 2011; Stallings et al., 1998).

2.7.3 Ditransitive sentences: Cross-linguistic comparison

With regard to the number of errors committed across languages (see the *Miscellaneous* row in Table 3.12), native speakers of Basque and Polish produced fewer overall errors than Spanish speakers ($\hat{\beta}$ = -1.458, z = -2.357, p = .02 and $\hat{\beta}$ = -1.218, z = -2.113, p = .04 respectively). Differences in error rates across languages might be due to the fact that verbs were presented in infinitive in the Spanish version of the experiment, and conjugated in Polish and Basque (see Section 2.3). There was no interaction between dependency length and language.

I now turn to cross-linguistic preferences for short dependencies in ditransitive sentences. Note that no analysis was feasible in the case of transitive sentences due to data sparsity. When long constituents were involved, there was a preference for producing a higher number of orders where heads are adjacent to each other, reducing dependency length, both when the IO is long ($\hat{\beta}$ = -1.626, SE = .250, z = -6.494, p < .001) – where canonical orders are those that yield shorter dependencies – and when the O is long ($\hat{\beta}$ = 1.866, z = 13.666, p <.001) – where non-canonical shifted orders yield shorter dependencies. When the IO was long as compared to short, Basque speakers produced canonical orders that yielded reduced dependency length less often than Polish speakers ($\hat{\beta}$ = 1.422, z = 5.203, p <.001). This interaction was also significant for Spanish versus Polish speakers, and the effect went in the same direction: compared to Polish speakers, Spanish speakers produced less canonical dependency length minimization-abiding orders than Polish speakers when the IO was long than when it was short ($\hat{\beta}$ = 1.955, z = 5.053, p <.001). This last interaction, however, needs to be interpreted cautiously. Unlike Basque and Polish speakers, who produced a high number of shifted orders when all constituents were short (41.3% and 36.5% respectively), Spanish speakers scarcely shifted constituents in the All-Short condition (10.1%), leaving very little room for any effect to be detected in this language when the IO was long versus when short (see Table 3.11) (note that to be able to find any effect of dependency length in this case in Spanish, the number of shifted word order in the Long-IO condition needs to be reduced in comparison the All-Short condition). Furthermore, there was an interaction between NP Position on the Screen and Length: When the to-be-arranged constituents were positioned on the screen in a left-to-right linearization consistent with the linearization in shifted sentences (e.g., O-IO in Basque and IO-O in Polish and Spanish), participants produced more shifted word orders when both constituents were short as compared to when IO was long ($\hat{\beta}$ = .718, z = 2.884, p <.004), and so obscured any possible effect of dependency length in the Long-IO condition for any of the three languages tested. Crucially, in the critical condition (Long-O), where there was now more room in Spanish for a difference in the strength of the effect of Length to be detected and no interaction between Length and NP Position was found, Basque speakers still produced significantly less orders that reduced dependency length as compared to Polish speakers ($\hat{\beta}$ = -1.466, z = -9.800, p <.001), but Spanish speakers did not differ from Polish speakers ($\hat{\beta}$ = -.130, z = -.652, p = .514). The results of the statistical analysis are reported in Table 3.15.

A clear pattern of results emerge from Experiment 2. Polish and Spanish, both head-initial languages (with S-V-O-(IO) canonical orders) but rich and poor respectively regarding case morphology, pattern in a similar way with regards to their preference for DLM, and both differ equally from Basque, a morphologically rich and head-final language (with S-(IO)-O-V canonical orders). Also, the preference for short dependencies is determined, at least in head-initial languages, by the type of construction at play: no effect of dependency length is found in transitive sentences in either Polish nor Spanish, whereas a quite robust effect emerges in ditransitive sentences in both languages. In Basque, on the contrary, the rather weak effect of DLM in transitive sentences is quantitatively similar to that found in ditransitive sentences.

	Estimate	SE	z-value	p-value
Ditransitive sentences (N = 1872)				
Intercept	-.590	.281	-2.099	< .04 *
Language: Spanish	-.573	.422	-1.358	.175
Language: Basque	.279	.352	.791	.429
Length: Long-IO	-1.627	.250	-6.496	< .001 ***
Length: Long-O	1.866	.137	13.666	< .001 ***
NP Position	-.439	.194	-2.267	< .03 *
Spanish * Long-IO	1.955	.387	5.053	< .001 ***
Basque * Long-IO	1.422	.273	5.203	< .001 ***
Spanish * Long-O	-.130	.200	-.652	.514
Basque * Long-O	-1.466	.150	-9.800	< .001 ***
Spanish * NP Position	-.419	.294	-1.428	.153
Basque * NP Position	-.0911	.214	-.426	.700
Long-IO * NP Position	.718	.249	2.884	<.004 **
Long-O * NP Position	.114	.130	.877	.380
Spanish * Long-IO * NP Position	.021	.386	0.055	.957
Basque * Long-IO * NP Position	-.517	.272	-1.900	.058 .
Spanish * Long-O * NP Position	-.200	.191	-1.047	.295
Basque * Long-O * NP Position	-.059	.144	-0.409	.682

TABLE 3.15: **Logit mixed model analysis for ditransitive sentences in Basque, Polish and Spanish in Experiment 2**

3 Discussion

In this chapter, a cued-recall production experiment was conducted to measure the strength of the effect of the preference for DLM in transitive and ditransitive sentences in three typologically diverse languages, Basque, Polish and Spanish. Specifically, I addressed the conditions in which the preference for shortening dependency length in production is to be found, modulated and canceled, which can in turn inform us about its cognitive underpinnings. Cross-linguistic experimental research grants us the possibility to control for sources of variation that were confounded in previous corpus research, namely rich case morphology and word order, and so it provides us with the opportunity to corroborate whether previous differences were the result of noise in the data due to any of the reasons outlined in the Introduction (correlated linguistic features, differences in corpus annotation schemes, on the type of texts used for the corpora, etc., see Section 1).

These results provide a more nuanced picture of the bias for short dependencies as a universal principle shaping language form. My main finding is that the general bias for reducing dependency length does not vary randomly. Instead, it is strongly modulated by a language-specific syntactic property such as word order. I first experimentally corroborated that, when presented with more than one alternative to convey the same propositional meaning, speakers of head-initial languages other than English, namely Spanish and Polish, also favor a *short-before-long* order that yields the shortest distance across phrasal heads. This is unlike verb-final languages, which favor the opposite *long-before-short* order. I thus confirm that basic word order affects the direction of linearization preferences (either *short-before-long* or *long-before-short*), which are claimed to emerge from a general preference for shortening dependency length. This supports the hypothesis that certain structures are preferred because they respond to general processing principles in the brain (Christiansen and Chater, 2008; Hawkins, 1994). Crucially, for the first time to the best of my knowledge, I demonstrate that word order determines the magnitude of the bias for reducing dependency length. In transitive and ditransitive sentences, Polish and Spanish, both head-initial languages but rich and poor respectively regarding case morphology, pattern similarly and radically differ from Basque, a morphologically rich head-final language. I argued that the modulation of the magnitude of the bias for DLM is due to differences in the strength of the expectations for the syntactic category to be found after a long dependency, which are in turn derived from distributional differences across languages' most frequent word orders.

Furthermore, the preference for short dependencies varies not only across but also within languages, even when possible sources of noise are controlled for. The strength of the tendency to reduce dependency length is highly dependent on sentence type. While there is a strong effect of dependency length in ditransitive sentences in both Polish and Spanish, I fail to find any effect in transitive sentences in these two languages (see similar results, where no effect is found, in English preverbal adjuncts and adverbs in Rajkumar, van Schijndel, White, and Schuler, 2016 and Temperley, 2007). This issue is addressed further below, in Section 3.3.

3.1 The role of word order based-expectations

As discussed in Section 1, differences in the strength of the effect of DLM might be accounted for in the framework of information-theoretic approaches of language production. Speakers will try to generate a robust signal that increases the probability of successful recovery of the to-be-conveyed message[6] (Fedzechkina et al., 2012; Futrell et al., 2015b; Gibson et al., 2013; Hawkins, 2014; Jaeger, 2010, etc.). Processing difficulty in long dependencies has been explained in terms of temporal decay and of interference among lexical items or discourse referents simultaneously activated in memory (Gibson, 1998; Grodner and Gibson, 2005; Lewis et al., 2006; Scontras et al., 2015; Van Dyke and Lewis, 2003, but see MacDonald et al., 2016). Speakers will tend to decrease processing difficulty by reducing dependency length. One available solution to do so is to rely on non-canonical shifted orders that minimize the distance across phrasal heads, so that syntactic and semantic relations are kept as local as possible, and information can be rapidly and easily recovered (Hawkins, 1994, 2014). Simultaneously, speakers will try to minimize the production cost derived from employing non-canonical shifted orders whenever possible[7]. I argued that either rich case making morphology or word order could help alleviate processing cost, and reduce the need to resort to DLM.

The results of Experiment 2 suggest that the probability of encountering a given linguistic category as measured by its cloze value affects the preference for shortening dependency length in production. For ditransitive sentences, as predicted, Basque shows a weaker preference for reducing dependency length than either Polish or Spanish, and these two do not differ from each other. Although in transitive sentences no statistical analysis was possible, my results suggest that, in this case, Basque tends to shorten dependency length more often than Polish or Spanish, as expected. This pattern of results complies with the hypothesis that cross-linguistic differences on the preference for short dependencies are due to distributional properties that stem from specific grammatical properties (in this case, linear word order) and that allow speakers and listeners to build expectations about upcoming lexical frames and linguistic elements, modulating the effect of memory-driven locality constrains. Until relatively recently, it was commonly assumed that the processing cost of a dependency increases monotonically as a function of its length, independently of the type of construction or the language in question. There is, however, growing evidence that distributional properties stemming from grammatical and lexical conventions allow speakers

[6]This idea appear to be supported by word order preferences across the three languages tested in ditransitive sentences when all constituents are short, in line with Fedzechkina et al. (2013). Polish and Basque, where case marking is present, presented greater word order flexibility than Spanish, with no case marking. In the case of Polish, however, this flexibility is conditioned by the order in which constituents appeared on the screen relative to each other.

[7]The claim that shifted orders are more costly than canonical ones is compatible with the miscellaneous data and errors. Both were more common in Polish and Spanish, with a higher rate of shifting, than in Basque, although differences across materials hinder any strong interpretation of these differences, since the greater number of errors in Spanish might be related to the fact that the main verb was presented as an infinitive in this language, while it was conjugated in Basque and Polish materials, which could have added extra difficulty to the task.

and listeners to build expectations about upcoming linguistic elements based on their Shannon information value in context (see Section 1.3). In fact, recent research indicates that both locality and surprisal effects orthogonally contribute to processing in languages such as English (Demberg and Keller, 2008), German (Boston, Hale, Kliegl, Patil, and Vasishth, 2008) or Hindi (Husain et al., 2014). One proposal argues that prediction strength influences the relation between locality and expectation effects (Husain et al., 2014). The strong predictability of a noun phrase has been shown to neutralize locality effects in a series of comprehension studies in Persian (Husain et al., 2014). When prediction strength is great, expectation effects dominate over locality effects. When prediction effects are weak, the relation is reversed, and locality effects emerge (but see Safavi et al., 2016, where no evidence supporting Husain et al., 2014 and colleagues' claim, but that argues for entropy reduction as a possible alternative measure). Under the assumption that language is designed for communicative efficiency (see Section 1), I outlined the hypothesis that knowledge of language-specific patterns could impact the way syntactic configurations tend to minimize dependency length also in production, given that the processing cost of long dependencies could be mitigated through expectation-based mechanisms.

Language production, as language comprehension, is sensitive to probabilistic distributions. In an elicited production study on the impact of constituent length on word order preferences in English, Stallings et al. (1998) found that argument structure biases impact production processes, mimicking for production the contextual effects of relative frequency profusely attested in comprehension (MacDonald, 1994; MacDonald et al., 1994; Trueswell et al., 1993, i. a.). Verbs that tend to appear separated from their complements, compared to those that need to appear adjacent to them, show slower elicitation latencies and undergo heavy NP shift more frequently, which renders dependency lengths shorter. Furthermore, speakers' productions have been shown to be sensitive to other types of language-specific probability distributions (Aylett and Turk, 2004; Bell et al., 2003; Gahl and Garnsey, 2004; Jaeger, 2010). In priming experiments, structural preferences are affected by recent experience with certain distributions, possibly through some form of implicit learning (Chang, Dell, and Bock, 2006) or adaptation expectation (Jaeger and Snider, 2013). Also, expectation-based mechanisms have been shown to play a role in language production in previous research. Gildea and Jaeger (2015), using syntactically annotated corpora, found that the sample of natural languages tested in their study (Arabic, Czech, German, English and Mandarin Chinese) showed a lower information density than expected by chance. In other words, those five languages tended to minimize (trigram) surprisal of words. Moreover, although dependency length and surprisal values were positively correlated, there was a trade-off with regard to optimizing them. According to the authors, this points to an independent influence of these two variables in language production. Both expectation and locality are required in current models of language comprehension and production, and further research is needed to address the issue of how locality and expectation interact with each other in language comprehension and production and whether their influence varies cross-linguistically.

3.2 The role of rich case marking

It has been argued that in morphologically rich languages, case marking can add redundancy to the signal by clearly conveying argument structure, reducing the processing cost of long dependencies and mitigating the need to resort to shifted orders (Futrell, 2017; Tily, 2010). The results from Experiment 2, however, lead us to discard the possibility that the coordinated activity of rich inflectional morphology and a universal preference for robust and efficient information transmission might result in a higher tolerance for long dependencies. I do not find morphological richness to be linked to a weaker preference for short dependencies. Polish patterns similarly to Spanish, a language with scant case morphology, and differs from Basque, a language, as Polish itself, with rich case morphology. My results are in line with recent comprehension data from Russian, a head-initial language with rich case-marking genetically related to Polish, where slower reading times were found at the verb by increased dependency distance, suggesting that case marking does not necessarily ameliorate processing difficulty, at least in certain circumstances (Levy and Keller, 2013). In any case, further research is necessary, that can tear apart some of the factors that might have influenced my results, especially the role of discourse considerations (see Section 3.4).

I am making the simplifying assumption that availability of case marking is the key factor, so that narrowing down the possible intended parsing even if only partially or probabilistically would facilitate processing. It is plausible that processing ease is dependent not only on cue availability (how often a given cue is present) but also on its reliability (given that a certain cue is present, how often it signals the same and only meaning). In Polish, cue reliability is possibly stronger than in Basque. Although Polish has a greater number of possible case endings that a lexical item might take, this variability is systematically conditioned by the target lexical item's phonological, semantic or lexical features. In Basque, on other hand, the agent (ergative singular) and the plural patient (absolutive plural) carry the same case suffix (*-ak*), so that speakers cannot rely on case marking alone to infer the intended meaning, and need to resort to other strategies such as word order or overall frequency of one or other case marking. Although in the transitive sentences in Basque plural agents were employed to avoid this ambiguity, it might be that it is overall redundancy in the system what is taken into account when mapping form and grammatical function. In this case, case marking would add a less degree of redundancy to the signal in Basque than in Polish. It would thus follow that Polish should show a weaker rate of DLM than Basque. My results, however, show exactly the opposite pattern. Additionally, cue reliability might be conditional on the linguistic context where the relevant cue is employed. That is, in Spanish, the preposition *a*, which can indicate either accusative, dative, allative, etc, should not be less reliable than case marking in any of the other two languages tested if the speaker/listener is aware of the lexical frame s/he is dealing with at a given moment. If so, even highly syncretic case marking could be reliable if it happens in complementary distributions. In this case, shifting rates across Basque, Polish and Spanish should not have differed from each other, which is also not was found in Experiment 2.

3.3 Differences across structures

In ditransitive sentences, all three languages tested in this study favored orders that minimized dependency length. However, in transitive sentences, Basque shows a preference for short dependencies that I fail to obtain in either Spanish or Polish. In previous experiments using a paradigm identical to the one employed in Experiment 2, other head-final languages aside from Basque also show an effect in transitive sentences. In Japanese and Korean, the proportion of shifted *long-before-short* orders in transitive sentences when the O was long reached 26% and 28% respectively (Dennison, 2008; Yamashita and Chang, 2001). In Basque, the proportion of shifted orders produced by participants in both transitive and ditransitive sentences reaches approximately 20%. In Korean, no ditransitive sentences were used. Although these differences could be related to the differences in prediction strength across types of constructions in the same language, there are other explanations that can account for them (further research is needed to test this possiblity). They are addressed below and in the next section.

Apart from considerations of efficient communication, it has been extensively shown that other factors influence the relative ordering of constituents, such as ambiguity, information status,accessibility, etc. (Bresnan et al., 2007; Wasow, 2002, i. a.). Differences across languages and linguistic varieties might emerge from differences in the hierarchical importance of each of these factors (Bresnan and Ford, 2010). For example, the preference for subject-initial orders is quite strong across languages in the world (Comrie, 1989; Greenberg, 1963; Tomlin, 1986). This can be a crucial factor in the restricted effect of dependency length in transitive sentences. Speakers may be reticent to place the S in a non-initial position, even when this leads to shorter dependency lengths. However, in Basque, an ergative language, subjects can be expressed by either ergative-marked or absolutive-marked NPs. This could lead to a greater tolerance for sentences beginning with absolutive-marked NPs that act as objects, due to their morphological overlap with absolutive-marked subject NPs that tend to appear in this position. Another possibility for the discrepancy between ditransitive and transitive sentences in Spanish and Polish is related to the impact of discourse considerations on word order, which I address in Section 3.4. In any case, further research is needed to assure that the effect of type of structure on the preference for short dependencies is indeed modulated by certain typological features

In previous corpus studies, a weaker effect of the bias for DLM has been reported for English preverbal adjuncts and adverbs (Rajkumar et al., 2016 and Temperley, 2007). The *noisy-context surprisal* hypothesis can account for differences in monotonicity in DLM between adjuncts and arguments (Futrell, 2017). According to Futrell (2017), arguments with a high amount of mutual information would tend to co-occur (see also Gildea and Jaeger (2015) and as far back as Behaghel (1909/10) or Bolinger (1957) and Givón (1991) for related claims, and Bybee (1985), who extends the so-called *principle of proximity* to morphological relations). Under this view, the preference for local dependencies is subsumed under a preference for

information locality. High predictable (that is, mutually highly informative) items will be processed more easily only if they are close to each other, because, given a noisy context, the possibility that some previous element is deleted or substituted will increase exponentially as a function of the distance between them. Consequently, elements in a given dependency with a weaker mutual information rate will not tend to be closer together since this would carry no processing benefit. Although a *noisy-context surprisal* account straightforwardly predicts the differences reported in the rate of DLM between adjuncts and arguments, it cannot explain the cross-linguistic paterrn of results reported in Experiment 2, because the amount of mutual information between arguments in transitive and ditransitive sentences is arguably identical across languages. Further assumptions about the distribution of noise depending on idiosyncratic properties of specific grammars would need to be included in the model proposed by Futrell (2017) to account for the reported differences across constructions in Experiment 3.

3.4 Further research: The role of discourse information

Further work on DLM will have to explore the relation between length- and discourse-triggered word order preferences across typologically diverse languages. Word order preferences have been shown to result from the combined action of a number of factors (see Bresnan et al. (2007) and references therein), such as definiteness, discourse accessibility, relative dependency length, pronominality, structural paralellism, etc. Specifically, given information has been claimed to tend to precede new information cross-linguistically (E. V. Clark and Clark, 1979; Haviland and Clark, 1974). Given information is more accessible and can be retrieved earlier, what gives the speaker more time to produce less available material later (Arnold et al., 2000; Bock and Irwin, 1980). Apart from in English, a *given-before-new* preference has been found experimentally in a production study in Japanese, a head-final language (V. S. Ferreira and Yoshita, 2003). Arnold et al. (2000) investigated both length and givenness and have found that they both influence word order preferences, at least in heavy NP-shift structures containing the strings *bring....to..* and *take...into account* and in the dative alternation with the verb *give* (but see Kizach, 2012) about confounds in their experimental design, especially in the coding of new and given material). Also, in a corpus study of transitive sentences in Polish, Siewierska (1993) provides strong evidence of the independent role of information structure and length on word order preferences in transitive sentences.

Crucially, givenness and length are correlated with each other (Arnold et al., 2000; Wasow, 2002). Phrases conveying old information tend to precede phrases conveying new information and, relevant to the study presented in this chapter, short phrases tend to correlate with old

information and complex ones with new information[8] (e. g., pronouns that refer to referentially given entities tend to be short). Therefore, in Polish or Spanish, both head-initial languages, the effect of dependency length could be increased by its correlation with information status, because they are both harmonically aligned (see Aissen, 1999), that is, they both pull in the same direction, namely *short-before-long* and *given-before-new*. In a head-final language such as Basque, however, the effect of short phrases and given information pull in opposite directions, namely *long-before-short* and *given-before-new*, cancelling each other out. This can account for the weaker effect of DLM in Basque ditransitive sentences when compared to Polish or Spanish. In Basque, the preference for *long-before-short* orders due to DLM would be reduced by its inverse correlation with the preference for *given-before-new* orders due to information packaging mechanisms. In Spanish and Polish, on the contrary, there is a positive correlation between both preferences, which could translate in an additive effect of both efficiency and discourse-related considerations. Also, the role discourse information in modulating word order preferences could explain some of the differences obtained in transitive and ditransitive sentences (see Section 3.3). In Basque, the long NP which had to be shifted before shorter ones was the O. In Spanish and Polish, the long NP which had to be shifted after the shorter ones was the S. Cross-linguistically objects are commonly associated with focused information, whereas subjects tend to act more commonly as topics or given information (Lee, 2015). The correlation between length (*short-long*) and discourse information (*topic-focus*) might thus not be as strong in transitive sentences as in ditransitive sentences, due to the syntactic position of the relevant phrase in transitive sentences. It is possible that this had led to a reduction of the preference for shifted word orders in Spanish and Polish in the critical Long-S condition in transitive sentences.

The information status of a given element can be marked through means of word order, intonation and/or morphological markers (for example, *ga*/*wa* in Japanese or *ka*/*nun* in Korean). Languages differ in how often they deploy these information packaging mechanisms. This could be a plausible explanation for a puzzling fact reported in previous work (Futrell, 2017), namely that the degree of word order freedom a language has can impact the strength of speakers' reliance on DLM strategies. Seemingly surprisingly, those languages whose grammars show a greater freedom to minimize dependency length seem to actually not be using it much (Futrell, 2017). So-called *scrambling* or *free word order* languages as Basque tend rely more on word order to convey information structure than fixed word order languages as English, which tend to use acoustic measures such as high pitch (Lambrecht, 1994). In languages which tend to use morphological devices or acoustic means rather than word order to mark information status, word order can still be employed to signal underlying syntactic structure, enhancing efficient processing. As word order is recruited to convey

[8]In Experiment 1 in Chapter 2, NPs in Basque tended to be placed in a preverbal position more often when long than when short. This indirectly supports the claim that complex NPs are often considered as new, non-presupposed information, since focused information in Basque needs to appear in a preverbal position (Altube, 1929; Hualde and Ortiz de Urbina, 2003; Ortiz de Urbina, 1989). In fact, marking focus prosodically in situ renders a sentence ungrammatical: *Liburua MIRENI eman dio* (*(S/he) book-the MARY-to gave*) versus **MIRENI liburua eman dio* (*(S/he) MARY-to book-the gave*)

information status, its reliability as a processing cue will decrease. This means that discourse considerations could influence word order more strongly in languages with a greater degree of word order freedom. In fact, the ranking of length and discourse effects might vary cross-linguistically (Wasow, 2002). Whereas in Polish discourse considerations are found to be is a stronger predictor of word order than length (Siewierska, 1993), the reversed situation is reported for English, a language with more strongly fixed word order (Arnold et al., 2000; Wasow, 2002). This difference, however, needs to be interpreted cautiously, since different experimental procedures and constructions were used in these two studies. In any case, word order freedom could be affecting the interplay of information and processing factors in determining word order alternations, distorting the effect of dependency length obtained in corpus and experimental studies so far. Further cross-linguistic experimental or corpus research where the possible impact of the correlation between information structure and dependency length in word order alternations is controlled for is thus highly relevant to investigate how strong the effect of DLM is, how widely efficiency thresholds vary across languages, and what implications this variation might have on the stability of a linguistic system and on language change.

3.5 Corpus and experimental studies: a fruitful relation

In this section, I will briefly address the issue of how my experimental results bear on previous corpus findings. Previous results in corpus studies reflect not only languages' grammars but also usage (see Temperley, 2018), where a rough estimate of the relative contribution of grammar and of syntactic choice to DLM in English is given). In this dissertations, only usage is investigated, specifically the preference for a given structural option when more than a syntactic choice is available for the same truth-conditional meaning. However, it is important to keep in mind that in all the corpus works cited in this chapter, where an above chance preference for short dependencies in languages' grammars is uncovered, researchers do implicitly or explicitly accept the hypothesis that this key feature is a grammatical property derived from a preference for efficient communication and robust information transmission (Fedzechkina et al., 2018; Futrell et al., 2015b; Haspelmath, 2008, *inter alia*). In other words, the existence of licensed word orders that yield shorter dependency lengths than their unlicensed counterparts is assumed to originate in previous usage preferences. Hawkins' Performance-Grammar Correspondence Hypothesis makes this claim explicit: Languages have conventionalized syntactic structures in proportion to their degree of preference in performance. I believe it is thus of crucial importance to explore whether the structures that are claimed to be more efficient are in fact preferred in on-line usage (see Chapter 2) and whether this preference varies across languages and why (see above, Chapter 3), which would allow us to investigate how well the language production system is attuned to communicate efficiently and what are the limits for this behaviour.

Lastly, I will argue that it is unclear whether the general preference for short dependencies in corpus studies stems, at least in some languages, from strategies other than DLM (this could be especially the case for the results in Liu, 2008). Head-final languages have been reported to use sentences where fewer arguments are present. For example, Japanese and Turkish (head-final) tend to use intransitive sentences more often than English and Spanish (head-initial) (Ueno and Polinsky, 2009), thus reducing the overall number of dependents to process in a sentence. This is claimed to be so because head-final languages are harder to process since more syntactic elements need to be kept in memory until the verb is reached (Ueno and Polinsky, 2009). If this were true, it would necessarily lead to a reduction in the overall dependency lengths calculated for Japanese or Turkish versus English or Spanish. However, the results reported by Ueno and Polinsky (2009) have not been replicated for all languages studied. Pastor and Laka (2013) find a similar percentage of transitive (with either 2 or 3 arguments) and intransitive sentences in both Basque and Spanish. In a corpus study of 2800 sentences in Spanish, they find that 50.8% sentences are intransitive in the Spanish corpus, and 52% in the Basque corpus. Further and cautious cross-linguistic corpus work is needed to address this issue and to investigate how other processing constraints might indirectly impact the preference for short dependencies.

4 Concluding remarks

By extending the research on the preference for DLM to typologically diverse languages in an experimentally controlled study, I present crucial evidence towards a better understanding of how language typology interacts with the cognitive bias towards short dependencies. The general preference for short dependencies, a fundamental principle in language comprehension, has been shown to also affect language production, having an impact both on grammars and usage preferences. The congruence between comprehension and production data is taken as evidence supporting the claim that the role of DLM in shaping language form is due to pressures for robust information transfer: Languages are highly efficient systems for transmitting information, adapted to our communication needs. An above-chance preference for short dependencies in a large number of languages is seen as proof of this efficient design.

In the study reported in this chapter, I show that a key typological property such as word order constraints the effect of DLM. Cross-linguistic variation in the preference for reducing dependency length is not random, but can be accounted from a rational-information theoretic approach to language processing. My results suggest that the effect of word order is mediated through how language-specific distributional information modulates expectations. So far, it has been showed that experimental evidence alongside data from corpus studies can fine-tune our hypotheses, open up new research questions and provide valuable information about the general mechanisms and representations behind what is believed to be a general cognitive principle and its impact on language form. Further cross-linguistic research in this topic needs

to address the extent of the impact of efficient communication on language form, the upper and lower bounds of DLM and its interaction with information packaging mechanisms and with other typological factors aside from head position and case marking.

5 Summary of findings

1. I find evidence supporting the claim that the strength of the bias for reducing dependency length in production is affected by word order-based expectations and not by case marking: In transitive and ditransitive sentences, Polish and Spanish, both head-initial but with a rich and poor case marking system respectively, pattern similarly and differ from Basque, a head-final language with a rich case-marking system.

2. The strength of the preference for short dependencies in production varies not only *across* but also *within* languages, even when possible sources of noise are controlled for. Unlike in head-final languages such as Basque, Japanese or Korean, transitive sentences in Spanish and Polish are unaffected by manipulations of dependency length.

3. Further research is need in order to investigate the interaction of dependency length minimization with other factors such as discourse considerations and degree of word order freedom.

Part II

Dependency length minimization across speakers

Chapter 4

Decreased efficiency in non-native language use

Abstract

Research on bilingual speakers' word order preferences can help us gain insight into the underlying nature of efficient considerations in language use, specifically in relation to the cross-linguistic bias for short dependencies. To this end, I investigated whether very proficient non-native bilinguals, who have acquired their non-dominant language very early differ from native speakers with respect to their preference towards dependency length minimization (DLM). Specifically, I analyzed word order preferences in non-native speakers of Spanish, whose dominant language was Basque (Experiment 3), and in non-native speakers of Basque (Experiment 4), whose dominant language was Spanish (Experiment 4), and compared them to native speakers' preferences. Spanish and Basque differ in two key ways. First, they present mirror-like *short-before-long* (Spanish) and *long-before-short* (Basque) linearizations when ordering constituents of varying length (see Chapter 2). Second, the effect of DLM is strong in Spanish but weak in Basque (see Chapter 3). Furthermore, efficiency in language use has been claimed to result from a balanced trade-off between reduction of signal redundancy and production ease, which could be tilted by non-native speakers towards the latter. The aforementioned experiments bear on cross-linguistic transferability of sentence production mechanisms such as DLM in bilingual speakers. In the case of DLM, transfer effects could be due to either superficial-based transfer of the linear direction of preferred word orders, to principle-based abstract transfer from the dominant to the non-dominant language or to a weaker effect of efficiency-related considerations. I found that, even in highly proficient early bilinguals, the tendency to reduce dependency length varied based on language dominance. Also, I do not find consistent evidence for superficial transfer of preferred word order linearizations. Although a subset of our results can be accounted for by the role of principle-based transfer (Experiment 3), the overall pattern of results that emerges from Experiments 3 and 4 and in previous literature (Dennison, 2008) is more in line with the hypothesis of a less efficient use of DLM by bilingual non-native speakers, which favor

production ease in detriment of robust information transmission. Further work on bilinguals with different levels of exposure to their non-native language can shed light on the interplay between transfer effects and efficiency considerations, as well as on their developmental trajectory through the lifespan of the bilingual speakers.

RESEARCH QUESTIONS:

- Is there superficial word order transfer effects in bilinguals of languages with opposing preferences regarding the linearization of constituents of different length, namely *short-before-long* (Spanish) and *long-before-short* (Basque)? That is, do bilinguals use in their L2 the length-triggered linearization favored in their L1?
- Is there abstract transfer effects in bilinguals whose native languages vary with regards to the magnitude of the preference for reducing dependency length, namely strong in Spanish and weak in Basque? That is, do bilinguals transfer the degree of reliance on DLM from their L1 to their L2?
- Do non-native speakers tip the balance between robust information transmission and production ease in favour of the latter, producing long dependencies more often than native speakers?

1 Introduction

In Experiments 3 and 4 in this chapter, I investigated the preference for reducing dependency length in high proficient non-native speakers of Spanish (VO) and Basque (OV) (AoA = 0-3 years). VO languages such as Spanish or English tend to place short elements before long ones. OV languages such as Basque or Japanese, on the contrary, tend to prefer the inverse ordering. These two opposite *short-before-long* and *long-before-short* word order preferences are claimed to emerge from the same universal cognitive bias to reduce the distance across phrasal heads (Futrell et al., 2015b; Hawkins, 2014, *inter alia*). Information-theoretic approaches to language argue that this bias is the result of efficient communication (Fedzechkina et al., 2018; Futrell et al., 2015b; Gildea and Jaeger, 2015; Gildea and Temperley, 2010): Speakers tend to produce short dependencies, which a large number of studies has been shown to be more easily comprehended (Bartek et al., 2011; Gibson, 1998, 2000; Grodner and Gibson, 2005). Data from bilingual speakers of languages which present mirror-like preferences with regard to the relative ordering of long and short constituents can be very informative with regard to the cognitive mechanisms underlying the bias for dependency length minimization. Additionally, non-native word order preferences can shed light on the developmental trajectory of an efficiency-related processing strategy across speakers immersed in bilinguals settings from a very early age, which could in turn have relevant applications for current language planning and educational policies.

Bilingual speakers of Basque and Spanish can provide relevant evidence on the source of the *short-before-long* and the *long-before-short* preferences found in languages with opposite head direction (e. g., Spanish - VO - and Basque - OV). They could transfer their knowledge of the sequential statistics of their native language to their non-native one. The *short-before-long* or the *long-before-short* ordering in the bilinguals' dominant language, Spanish or Basque, could persist and override the most efficient ordering in the less dominant one. That is, L1Basque-L2Spanish speakers (hereafter, L1Ba-L2Sp) would show a *long-before-short* tendency when using either their L1 Basque or their L2 Spanish. Correspondingly, L1Spanish-L2Basque speakers (hereafter, L1Sp-L2Ba) would show a *short-before-long* preference in either their L1 Spanish or their L2 Basque. Any results in this line would cast doubts on the claim that *short-before-long* and *long-before-short* orders are the result of the same universal processing preference towards reducing dependency length. In this case, a further assumption would be needed to account for the emergence of superficial transfer effects of preferred linearizations, namely that switching from a *long-before-short* to a *short-before-long* linearization (and viceversa) takes a certain amount of experience with the non-native language to develop, since it requires increased familiarity with its head direction (for this argument, see Dennison, 2008). However, this assumption could be problematic: The amount of linguistic experience required to switch to the non-native linearization should be significantly large, since Dennison (2008) was also testing extremely proficient early bilinguals.

Another possibility is that native language word order preferences will influence non-native ones not through superficial but through abstract or principle-based transfer. This possibility is especially plausible if learning the syntax of a new language is based more strongly in abstract representations than in distributional knowledge (see Fedzechkina et al., 2018 for this argument). In fact, although both are shown to contribute to language learning, recent evidence indicates that, when pitted against each other, language learners privilege structural over superficial information (Culbertson and Adger, 2014). DLM would be the ultimate cause of opposing length-triggered linearizations, but the strength of this preference could be transfered by bilingual speakers from their native to their non-native language. There is evidence that the strength of certain cues or processing strategies in the native language impact processing strategies in the non-native one and that degree of exposure to a certain property is a key factor modulating non-native preferences (for a review, see Frenck-Mestre, 2005). In the case of non-native bilinguals, in ditransitive sentences, L2Sp-L1Ba speakers will show a weaker effect of DLM when using their L2 Spanish, whereas L2Ba-L1Sp speakers will present a stronger effect of DLM than natives, when speaking their L2 Basque. In contrast, in transitive sentences, the expected pattern would be the opposite. This is so because, as demonstrated in Chapter 3, in ditransitive sentences, Spanish shows a much stronger effect of DLM than Basque, while the pattern is reversed for transitive sentences.

Within the theoretical framework of this dissertation, there is a remainig plausible scenario that can predict differences across native and non-native speakers with regard to the preference for dependency length minimization. DLM is claimed to respond to efficient

considerations and to seek communicative success while minimizing production effort by means of reducing any redundancy in the linguistic signal (Fedzechkina et al., 2012; Gibson et al., 2013; Horn, 1984; Jaeger, 2010; Zipf, 1949, *inter alia*). It is possible that the trade-off between facilitation in production and the need to ensure a high probability of communicative success is not balanced in the case of non-native speakers. Non-native speakers would favour ease of production, due to the higher attentional and memory demands imposed by using a non-native language. As argued in Section 1.2 in Chapter 3, producing word orders with short dependencies might entail a greater production effort than producing their canonical counterparts, if all possible candidates compete for activation. Competition for selection among alternative structures restricts the availability of each choice, leading to increased production cost (Dell and O'Seaghdha, 1994; Hwang and Kaiser, 2014; Myachykov et al., 2013). Non-native speakers might face more difficulties than native-speakers in maintaining more than one competing order in memory, or in retrieving it from memory later on. This implies discarding the non-canonical order, since more frequent forms have higher base activation. In other words, the tendency to reuse more frequent sentence plans saved in long-term memory might be helpful when facing a demanding production task (MacDonald, 2013). This would entail that non-native speakers, either L2Sp-L1Ba or L2Ba-L1Sp, would show a stronger preference for canonical orders, especially when production cost is high, as it is assumed to be the case when long dependencies need to be used. This prediction is compatible with theories of first and second language acquisition and language contact that argue for a tendency towards over-regularization in language learners (Lupyan and Dale, 2010). Linguistic inconsistencies, especially when patterns are probabilistic in nature and not contextually dependent nor consistent across speakers, as is the case for Heavy NP shift (see individual differences across items and subjects in Chapters 2 and 3), tend to be regularized both in the case of children (Newport, 1999; Singleton and Newport, 2004) and adults (Hudson Kam and Newport, 2016).

There is one previous production study that compared length-triggered word order preference in bilingual speakers of languages with opposite linearization preferences, namely English (VO) (*short-before-long*) and Korean (OV) (*long-before-short*). Dennison (2008) investigated the preference for reducing dependency length in transitive sentences in Korean monolingual speakers (late learners of English), L1Korean-L2English bilinguals, Equi-dominant bilinguals and L1English-L2Korean bilinguals. Korean monolingual speakers and L1Korean-L2English bilinguals produced OSV orders 29.4% and 41.9% more often respectively when the O was long than when it was short[1] (note that when both constituents are short there is no processing advantage in shifting their order). In contrast, Equi-dominant and L1English-L2Korean bilinguals used mostly canonical SOV orders even when the O was

[1]Differences between Korean monolingual speakers and L1Korean-L2English speakers are reported as non significant. However, Dennison (2008) employed small and unequal sample sizes (11 participants in each of the three bilingual groups and 48 monolinguals of Korean), which can affect the homogeneity of variance assumption of ANOVAs. Furthermore, data was analyzed using ANOVAs, in spite of the dependent variable being binary (for the shortcomings of ANOVA analysis of categorical outcomes, see Jaeger, 2008, *inter alia*). This likely reduces the reliability of the statistical results reported in his experiment.

long, resorting very rarely to shifted *long-before-short* OSV orders (less than 7% of the time). These results were interpreted as evidence that the superficial *short-before-long* word order in the bilinguals' dominant language, English, could override the *long-before-short* order in the less dominant one, Korean. However, it cannot be discarded that any differences across groups are derived from the strength of the preference to reduce dependency length in English versus Korean transitive sentences. Although corpus studies suggest that English is a language specially attuned to the bias for DLM (Gildea and Temperley, 2010), it is quite plausible that it should present no effect of DLM in transitive sentences, like other head-initial languages such as Spanish and Polish (see Chapter 3), since, in transitive sentences, English allows permutations of sentence constituents in only very exceptional cases as when there is topicalization or left dislocation (Gregory and Michaelis, 2001; Prince, 1995). Speakers more exposed to English than to Korean could rely less on DLM in transitive sentences when speaking their less dominant language, Korean, due to the weaker degree of reliance on DLM in their most dominant one, English. Additionally, in head-final languages like Korean, SOV canonical order overlaps with a *short-before-long* order when the O is long, resulting in longer dependencies. In consequence, the stronger preference for canonical orders could be accounted for by a preference for production ease in detriment of robust information transfer in the case of non-native speakers. In conclusion, it is unclear whether any absence of shifting in Equi-dominant and L1English-L2Korean speakers in Dennison (2008) is due to superficial or abstract word order transfer from English or to their need to facilitate production over complying with language efficiency requirements.

Summing up, in this chapter, I investigated the magnitude and linearization direction (either *short-before-long* or *long-before-short*) of the bias for DLM in L2Sp-L1Ba (Experiment 3) and L2Ba-L1Sp (Experiment 4) high proficient bilinguals (AoA = 0-3 years) when speaking their non-native language. These two groups of bilinguals are key to address the issue of the universality of DLM as a processing principle and whether native speakers' preferences respond better to efficiency considerations when compared to very proficient non-native speakers. First, if superficial transfer takes place, L1Ba-L2Sp speakers should find a *long-before-short* preference in Spanish and L1Sp-L2Ba bilinguals, a *short-before-long* one in Basque, contrary to the results reported for native speakers of these languages (see Chapters 2 and 3). Second, if any influence of native on non-native production preferences is due to abstract principle-based transfer, then, in Spanish, L1Ba-L2Sp bilinguals should show a weaker preference for DLM in ditransitive sentences and a stronger one in transitive sentences. In contrast, in Basque, compared to native speakers, L1Sp-L2Ba bilinguals should show a stronger effect of dependency length in ditransitive sentences and a weaker one in transitive sentences. Last, if non-native speakers favor reducing production effort over robust information transmission, the preference for DLM should be weaker in all cases due to a stronger tendency to use canonical orders, and non-natives should produce inefficient long dependencies more often than natives in any L2.

Lastly, I also explored the effect of dependency length on word order preferences regarding

verb position in non-native speakers of Basque. In Chapter 2, I argued that, aside from general processing mechanisms such as adjacency of phrasal heads, language-specific resources could be engaged to enhance communicative efficiency. In verb-final flexible languages like Basque, verb agreement could be deployed with this aim. Unlike in Korean or Spanish, verb position is very flexible in Basque. This means that both verb-medial and verb-final orders are common in the language. Basque also presents pluripersonal agreement: the verb agrees with the subject, the object and the indirect object (see Section 1.1 in Chapter 2 for a brief grammatical sketch of Basque). When long elements are shifted relative to short ones, resulting in a less frequent non-canonical O-IO order, verb agreement could be used as a signaling device that would facilitate faster integration of upcoming verb arguments and easier recognition of the intended syntactic structure. In consequence, when shifted orders are used in Basque, there should be a stronger tendency to produce verb medial orders. This prediction is met in the case of native speakers of Basque (see Section 2.7.3 in Chapter 2). Now I explored L2Ba-L1Spa bilinguals' word order preferences with regard to the position of the verb in sentences where long dependencies are involved, and whether they behave like native speakers.

I will next list the predictions made by the three proposals outlined for non-native bilinguals' word order preferences with respect to verb position when using their L2. Predictions are less clear for word order preferences regarding verb position than regarding the relative ordering of short and long constituents. In consequence, any evidence in this regard should be treated cautiously. In Spanish, the verb commonly appears in the middle of the sentence, both in transitive and ditransitive sentences. First, if superficial transfer is an active force behind word order preferences, one could expect that, in any case, L1Sp-L2Ba speakers would rely more on verb-medial orders than L1Ba-L2Sp speakers, independently of dependency length. Second, if abstract transfer takes places, it is then possible that L2Ba-L1Sp bilinguals do not rely on a efficiency-enhancing mechanism that consists on displacing the verb to another position in the sentence when speaking Basque, since verb position in Spanish is quite rigid. In this case, one could argue that, if L2Ba-L1Sp speakers transfer the tendency to produce verbs in a fixed position, they should resort more to verb-final canonical orders in all conditions when using Basque. If differences across native and non-native speakers are the result of an unbalanced trade-off between reducing production effort and ensuring communicative success, L2Ba-L1Sp speakers should produce more canonical verb-final orders, especially when dependencies are long.

2 Experiment 3: Spanish native and non-native speakers

2.1 Overview of the experiments

This experiment was, as Experiment 1 and 2 in Chapters 2 and 3, a cued-recall production task where participants were presented with unordered phrases of different lengths on a computer

screen and were asked to construct sentences with them. Transitive and ditransitive sentences were presented. I investigated the preference for producing shifted orders as a function of constituent length in transitive and ditransitive sentences in Spanish. Here I was interested in whether the bias for DLM would vary across speakers as a function of their exposure to languages that employ opposing orders to obtain reduced dependency lengths (*short-before-long* in Spanish and *long-before-short* in Basque) and that differ with respect to the magnitude of this preference (strong in Spanish and weak in Basque). I tested extremely high proficient non-native speakers of Spanish, whose dominant language was Basque (L2Sp-L1Ba), and compared them with both monolinguals and bilingual native speakers of Spanish whose more dominant language was Spanish (L1Sp-L2Ba).

Let me review the predictions made by the set of hypotheses put forth in the previous section. I will focus here in what is expected for Spanish native and non-native speakers. If there is superficial transfer, L2Sp-L1Ba speakers will rely on a *long-before-short* strategy more often than monolinguals or L1Sp-L2Ba speakers, namely *long-before-short* O-V-S when O is long in transitive sentences and *long-before-short* S-V-IO-O when IO is long in ditransitive sentences. I address this issue in Experiment 4, where comparisons between critical conditions are feasible. If reliability or frequency of a certain production strategy in the native language, in this case DLM, impacts non-native routines, then, for ditransitive sentences, it would be expected that L2Sp-L1Ba speakers will resort to DLM less than L1Sp-L2Ba speakers, leading to the production of a fewer number of *short-before-long* S-V-IO-O orders when O is long compared to native speakers of Spanish. For transitive sentences, on the contrary, a stronger reliance on DLM should be found in L2Sp-L1Ba compared to native speakers. This means that L2Sp-L1Ba speakers should produce more shifted *short-before-long* O-V-S orders when the S is long. Last, if unbalanced trade-off between production effort and robust information transmission is the cause for any difference across speakers, L2Sp-L1Ba speakers will produce more canonical orders than L1Sp-L2Ba or monolingual speakers, especially when dependencies are long.

2.2 Participants

18 high proficient L2Sp-L1Ba bilinguals (13 females, mean age = 20.61, AoA = 5.44, SD = 2.58) were recruited and tested in Spanish, and compared to the results obtained from L1Sp-L2Ba bilinguals (15 females, mean age = 19.39) reported in Chapter 4. A group of 20 Spanish monolinguals was also included as a baseline (15 females, mean age = 20.61). 1 had to be excluded due to experimental error. Monolinguals were included in order to control for any possible effects that using Basque, even as L2, could have in word order preferences in an L1, due to its weaker preference for shortening dependency lengths compared to other languages tested (see Chapters 2 and 3). As in all previous experiments, language background was assessed using the questionnaire reported in Appendix 1. Our L2Sp-L1Ba participants reported to use Basque as their primary language for daily communication (see Table 4.4) and

rated their overall Spanish proficiency level as 5.53 over a maximum of 6 (SD = .97) (see Table 4.2). For comparisons with the group of L1Sp-L2Ba bilinguals from Experiment 2, see Tables 4.3 and 4.1 below. Monolinguals reported Spanish as their only vehicle of daily communication.

As in previous experiments, all our bilingual participants had obtained a proficiency certificate in Basque equivalent to a C1 level in the Common European Framework, and their education previous to their undergraduate studies was conducted exclusively in Basque. Our L2Sp-L1Ba participants came from areas with a high density of bilinguals: 5 came from areas with a bilingual population of above 75%; 12 came from cities with a density of bilinguals between 50-80%; and 1 came from an area with 20-50% bilingual speakers (IV Sociolinguistic Map, 2006; Censo de población y vivienda, 2001). On the contrary, from our L1Spanish-L2Basque participants, none came from cities with a density of bilinguals over 75%, 8 came from areas with population of bilinguals between 50-75%, and 10 from areas with less than 35% bilinguals. (IV Sociolinguistic Map, 2006; Censo de población y vivienda, 2001). All participants were naïve to the purpose of the experiment, filled in a consent form in Spanish and received an economic compensation for their participation.

	Basque	**Spanish**
Listening	6.72 (.46)	6.89 (.32)
Speaking	6.22 (.73)	6.66 (.48)
Reading	6.61 (.50)	6.83 (.38)
Writing	6.28 (.67)	6.61 (.50)
Mean	6.46 (.52)	6.75 (.33)

TABLE 4.1: **Proficiency level scores (SD in brackets) as reported by L1Sp-L2Ba bilinguals in Experiment 2**. *Self-assessed language proficiency is rated on a 7-point scale: 7 = Native-like level; 6 = High level; 5 = Medium-high level; 4 = Medium level; 3 = Medium-low level; 2 = Low level; 1 = No knowledge*

	Basque	**Spanish**
Listening	6.89 (.32)	6.17 (.86)
Speaking	6.83 (.38)	5.00 (1.28)
Reading	6.83 (.38)	5.72 (1.02)
Writing	6.61 (.61)	5.22 (1.26)
Mean	6.79 (.29)	5.53 (.97)

TABLE 4.2: **Proficiency level scores (SD in brackets) as reported by L2Sp-L1Ba bilinguals in Experiment 3**. *Self-assessed language proficiency is rated on a 7-point scale: 7 = Native-like level; 6 = High level; 5 = Medium-high level; 4 = Medium level; 3 = Medium-low level; 2 = Low level; 1 = No knowledge*

	Infancy (0-3 years)	**Primary school (4-12 years)**	**Secondary school (12-18 years)**	**Adulthood (after 18 years)**
Home	6.39 (.98)	6.56 (.78)	6.44 (.86)	6.28 (1.23)
School		2.56 (1.89)	2.78 (1.96)	4.28 (1.96)
Other		4.94 (1.11)	4.89 (1.13)	5.00 (1.08)
Equally comfortable in both languages: 7 ; More comfortable with Spanish: 11				

TABLE 4.3: **Language history (SD in brackets) as reported by L1Sp-L2Ba bilingual in Experiment 2**. *Self-assessed language use is rated on a 7-point scale:* 1 = Only Basque; 2 = Mostly Basque, a few times Spanish; 3 = Mostly Basque but Spanish at least 25% of the time; 4 = Basque and Spanish equally; 5 = Mostly Spanish but Basque at least 25% of the time; 6 = Mostly Spanish, a few times Basque; 7 = Only Spanish

	Infancy (0-3 years)	**Primary school (4-12 years)**	**Secondary school (12-18 years)**	**Adulthood (after 18 years)**
Home	1.61 (.70)	1.56 (.62)	1.67 (.77)	1.67 (.77)
School		1.67 (.69)	2.17 (1.34)	2.06 (1.26)
Other		1.83 (.71)	1.94 (.73)	2.33 (.91)
Equally comfortable in both languages: 2; *More comfortable with Basque:* 16				

TABLE 4.4: **Language history (SD in brackets) as reported by L2Sp-L1Ba in Experiment 3**. *Self-assessed language use is rated on a 7-point scale:* 1 = Only Basque; 2 = Mostly Basque, a few times Spanish; 3 = Mostly Basque but Spanish at least 25% of the time; 4 = Basque and Spanish equally; 5 = Mostly Spanish but Basque at least 25% of the time; 6 = Mostly Spanish, a few times Basque; 7 = Only Spanish

2.3 Materials and Procedure

Task, stimulus presentation and materials were identical to Experiments 1 and 2 (see Sections 2.3 and 2.4 in Chapter 2 and Sections 2.3 and 2.4 in Chapter 3). The experiment was conducted exclusively in Spanish.

2.4 Scoring

Scoring followed the same procedure as in Experiments 1 and 2 (see Sections 2.5 in Chapter 2 and Section 2.5 in Chapter 3). For Spanish monolinguals, a total of 16 (3.7%) transitive and 61 (11.3%) ditransitive sentences were coded as miscellaneous. For L2Sp-L1Ba bilingual speakers, a total of 23 (5.3%) transitive and 111 (20.6%) ditransitive sentences were coded as miscellaneous. L1Sp-L2Ba committed a total of 19 (4.4%) and 50 (9.3%) errors in transitive and

ditransitive sentences respectively, as reported in Section 2.5 of Chapter 3.. All miscellaneous data were discarded from further analysis.

Changes in tense were permitted, as well as synonym substitutions, and minor deletions, changes and additions were included in the main analysis. Our scoring resulted in a total of 416 (96.3%) transitive and 479 (88.7%) ditransitive target sentences for Spanish monolingual speakers and 409 (94.68%) transitive and 429 (79.44%) ditransitive target sentences for L2Spanish-L1Basque bilinguals. As reported in Section 2.5 in Chapter 3, L1Sp-L2Ba bilinguals produced a total of 413 (95.6%) transitive 490 (90.7%) ditransitive and target sentences.

2.5 Analysis

As in previous experiments, the maximum random effect structure justified by the data was used, using model comparison between all the models that converged and did not contain correlations between the random effects equal to 1/-1 (Baayen et al., 2008). I report p-values estimated by the R-package lmerTest (Kuznetsova et al., 2017). All analyses were conducted using R (R Core Team, 2014) and the lme4 package (D. M. Bates et al., 2014). As in Spanish in Experiment 2 (see Section 2.7.1), transitive sentences had to be excluded from any analysis due to data scarcity (see Section 4.5 in Chapter 3). For the ditransitive sentences, I fitted the data from Spanish monolinguals, L1Sp-L2Ba and L2Sp-L1Ba bilinguals to a model with Length, Group and NP Position on the Screen as well as their interaction as fixed effects, and by-subject and by-item random intercepts and by-subject slopes. As in previous experiments, Length was Helmert-coded and NP Position on the Screen was centered. I used the L2Sp-L1Ba bilinguals as reference group in this model. Two other models were fitted to data only from the monolingual and the L2Sp-L1Ba participants respectively, with Length and Position in the Screen and their interaction as fixed effects, and with by-subject and by-item random intercepts and by-subject slopes for Length. As in the previous model, NP Position on the Screen was centered and Length Helmert-coded.

2.6 Results

2.6.1 Transitive responses

As L1Sp-L2Ba speakers in Experiment 2 (see Section 2.7.1 in Chapter 3), monolinguals and L2Sp-L1Ba bilingual speakers produced virtually no shifted orders constituents and, therefore, no statistical analysis was conducted. Results are reported in Table 4.5

Group	**Length condition**		
	All-Short	Long-S	Long-O
Spanish monolinguals	0	0 (0%)	0 (0%)
Bilinguals L1Spanish-L2Basque	0 (0%)	0 (0%)	1 (.1%)
Bilinguals L2Spanish-L1Basque	0 (%)	1 (.2%)	0 (%)

TABLE 4.5: **Raw data and percentages (in brackets) of shifted responses in transitive sentences in Spanish monolinguals and bilingual L1Sp-L2Ba and L2Sp-L1Ba speakers in Experiment 3.** The experiment was conducted in Spanish.

2.6.2 Ditransitive responses

The critical group in this experiment, L2Sp-L1Ba bilingual speakers, produced more errors when a constituent was long than when short, both in the case of the object ($\hat{\beta}$ = .906, z = 3.017, p <.003) and of the indirect object ($\hat{\beta}$ = 1.015, z = 3.398, p <.0007). Monolinguals showed the same pattern (O: $\hat{\beta}$ = .869, z = 2.320, p = .02; IO: $\hat{\beta}$ = .794, z = 2.093, p = .04). L1Sp-L2Ba speakers showed numerically the same trend, but it did not reach significance (O = $\hat{\beta}$ = .479, z = 1.181, p = .2; IO: $\hat{\beta}$ = .609, z = 1.525, p = .1). When results where compared across groups, L2Sp-L1Ba speakers committed an overall higher rate of miscellaneous responses than both monolinguals and L1Sp-L2Ba speakers and this difference reached marginal significance in both cases ($\hat{\beta}$ = .774, z = 1.776, p = .08 and $\hat{\beta}$ = .726, z = 1.659, p = .097, respectively). Spanish monolinguals and L1Sp-L2Ba speakers did not differ from each other ($\hat{\beta}$ = .048, z = .102, p = .919). There was no interaction between native language and condition (all ps > .4). These results correspond to models with only by-subject and by-item random intercepts.

Results relative to length-triggered word order preferences from all the three groups that participated in Experiment 3 are reported in Tables 4.6 and 4.7 and graphically represented in Figure 4.1. NP Position on the Screen was kept as a variable in all tables and figures in this chapter. Although this predictor dit not interact with other variables in any of the models reported in this chapter, it did in the case of L1Sp-L2Ba bilinguals (see Table 3.14 in Chapter 3), and it contributed to explain away the variability in the current data set: The comparison across groups reported in Table 4.10 shows that all participants, independently of their L1, tended to use less shifted O-IO orders when this was not the linearization previously seen on the screen, that is, when the order of constituents presented on the screen was inconsistent (O-IO) with the expected direction of shift in Spanish (IO-O) ($\hat{\beta}$ = -.529, z = -2.529, p = .01). As expected, there was a general tendency to place the object later in the sentence more often when it was long than when it was short, yielding a *short-before-long* linearization preference across all groups ($\hat{\beta}$ = 1.318, z = 7,613, p < .0001). This preference also obtained when Spanish monolinguals and L2Sp-L1Ba bilinguals were considered separately ($\hat{\beta}$ = 2.364, z = 6.747, p < .0001 and $\hat{\beta}$ = 1.410, z = 5.816, p < .0001) (see Tables 4.8 and 4.9). This is especially relevant in the case of L2Sp-L1Ba

bilinguals, since it is unlike what they do in their native language, Basque, where the opposite *long-before-short* preference obtains (see Section 2.7.2 in Chapter 2).

In the model where the three groups of speakers were compared (see Table 4.10), an interaction emerged between Length and Group. Both Spanish monolingual speakers and L1Sp-L2Ba bilinguals showed a stronger effect of Length than L2Sp-L1Ba bilinguals ($\hat{\beta}$ = .884, z = 3.505, p < .001 and $\hat{\beta}$ = .651, z = 2.680, p < .008, respectively). Monolinguals and L1Sp-L2Ba speakers did not differ from each other in the model ($\hat{\beta}$ = .233, z = .885, p = .4, when reference level was changed to monolinguals). However, the tendency went numerically in the same direction as when monolinguals were compared to L2Sp-L1Ba participants: The proportion of shift when the O was long compared to when short was smaller for L1Sp-L2Ba bilinguals compared to Spanish monolinguals. When the IO was long, no meaningful comparison between the All-Short and Long-IO conditions was possible due to the floor effect in the former. Numerically, all groups of participants produced more *short-before-long* orders when the IO was long than when short, up to 6% in the case of both monolinguals and L2Sp-L1Ba.

Group	**Length condition**		
	All-Short	Long-IO	Long-O
Spanish monolinguals	13 (7%)	3 (1%)	134 (84%)
Bilinguals L1Spanish-L2Basque	17 (10%)	15 (9%)	138 (85%)
Bilinguals L2SpanishL1Basque	16 (10%)	6 (4%)	81 (59%)

TABLE 4.6: **Raw data and percentages (in brackets) of shifted responses in ditransitive sentences in Spanish monolinguals and bilingual L1Sp-L2Ba and L2Sp-L1Ba speakers in Experiment 3**. The experiment was conducted in Spanish.

Group	**NP Position on the Screen**	**Length condition**		
		All-Short	Long-IO	Long-O
Spanish monolinguals	*Inconsistent with shift direction*	1 (01%)	1 (01%)	64 (79%)
	Consistent with shift direction	12 (13%)	2 (02%)	74 (89%)
Miscellaneous		11	24	26
L1Spanish-L2Basque	*Inconsistent with shift direction*	1 (1%)	7 (9%)	60 (75%)
	Consistent with shift direction	16 (19%)	8 (10%)	78 (95%)
Miscellaneous		12	20	18
L2SpanishL1Basque	*Inconsistent with shift direction*	3 (.04%)	2 (.03%)	36 (53%)
	Consistent with shift direction	13 (16%)	4 (.06%)	45 (65%)
Miscellaneous		22	46	43

TABLE 4.7: **Raw data and percentages (in brackets) of shifted responses relative to position on the screen in Spanish monolinguals and bilingual L1Sp-L2Ba and L2Sp-L1Ba speakers in Experiment 3.**

	Estimate	**SE**	**z-value**	**p-value**
Ditransitive sentences (N = 508)				
Intercept	-2.117	.582	-3.640	< .001 **
Length: Long-IO	-.541	.721	-.751	.453
Length: Long-O	2.364	.350	6.747	< .0001 ***
NP Position: IO-O	-.785	.301	-2.604	< .0091 **
Length (Long-IO) * NP Position	.518	.420	1.233	.218
Length (Long-O) * NP Position	.072	.171	.423	.673

TABLE 4.8: **Logit mixed model analysis for ditransitive sentences in monolingual speakers of Spanish in Experiment 3**

	Estimate	**SE**	**z-value**	**p-value**
Ditransitive sentences (N = 429)				
Intercept	-2.404	.508	-4.734	< .001 ***
Length: Long-IO	-.390	.574	-.680	.497
Length: Long-O	1.410	.242	5.816	< .0001 ***
NP Position	-.560	.212	-2.639	< .008 **
Length (Long-IO) * NP Position	.256	.301	.850	.396
Length (Long-O) * NP Position	.145	.120	1.212	.226

TABLE 4.9: **Logit mixed model analysis for ditransitive sentences in L1Ba-L2Sp bilingual speakers of Spanish in Experiment 3**

Finally, to ensure that the effect of dependency length was not due to a small subset of items or subjects in any of the groups, the results by item and by subject were examined. Of the 54 subjects in the three groups, all of them showed a preference to utter the object later when it was long compared to when short (see Figure 4.2). In turn, when word order preferences across items were inspected, all groups showed a preference to place the object phrase after the indirect object phrase when the former was long more often than when both were short (see Figure 4.3). The only exceptions were items 9, 14, 22 and 29 in the group of L2Sp-L1Ba speakers. As for Basque in Chapter 2, there were numerical differences with respect to the magnitude of the effect of dependency length in both items and subjects. Differences across subjects will be covered in Chapter 5.

In this experiment, I replicated two main findings reported in Chapter 3. First, a *short-before-long* preference was obtained across all types of speakers of Spanish, both monolingual and bilinguals, independently of whether Spanish was their L1 or their L2. Second, as for native bilingual speakers in Chapter 3, the preference for reducing dependency length is dependent on type of sentence both for monolinguals and for non-native speakers. There was no preference for DLM in transitive sentences. Also, this experiment revealed a

series of new findings. L2Sp-L1Ba bilinguals shift long constituents after short ones when speaking in Spanish, contrary to what they do in their native language, where they prefer to shift long constituents before short ones (see Chapter 2). Crucially, L2Sp-L1Ba bilinguals showed a weaker tendency towards shifted *short-before-long* orders than native speakers of Spanish, both monolinguals and L1Sp-L2Ba bilinguals.

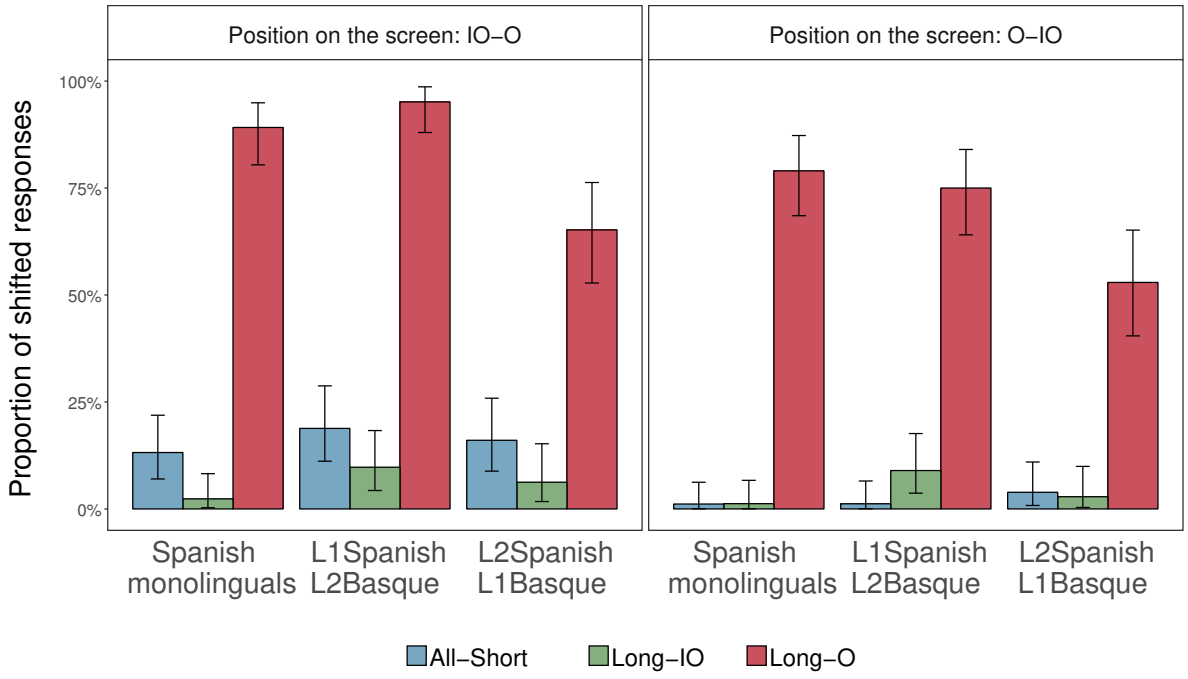


FIGURE 4.1: **Proportion of shifted responses in ditransitive sentences in Spanish monolinguals and bilingual L1Sp-L2Ba and L2Sp-L1Ba speakers in Experiment 3**. The error bars show 95% confidence intervals

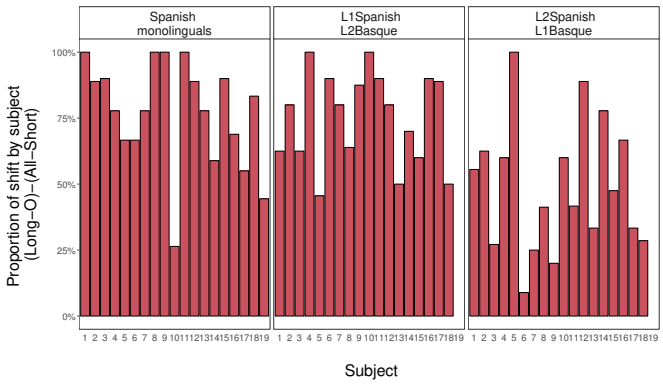


FIGURE 4.2: **Proportion of shifted responses by subjects in Experiment 3**. Positive proportions indicate a *short-before-long* ordering. No barplot indicates a preference for canonical orders. Negative proportions indicate a *long-before-short* ordering

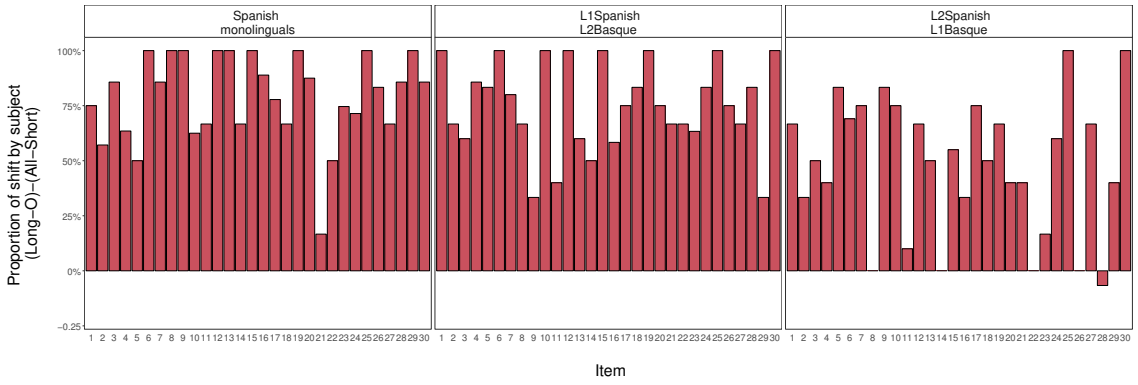


FIGURE 4.3: **Proportion of shifted responses by items in Experiment 3**, Positive proportions indicate a *short-before-long* ordering. No barplot indicates a preference for canonical orders. Negative proportions indicate a *long-before-short* ordering

2.7 Discussion

In Spanish transitive sentences, non-native speakers do no show any effect of dependency length, as was the case for native speakers (see Chapter 3), and unlike in their native language, Basque, where a *long-before-short* preference has been found (Ros et al., 2015 and Chapter 2). That is, when O is long, they do not produce shifted orders according to the *long-before-short* preference present in Basque, their native language. Also, when the S is long, they do no produce more shifted *short-before-long* OSV orders than native speakers, and neither do resort more often to DLM. In other words, in transitive sentences, I do not find evidence for the predictions from superficial or abstract transfer. However, the lack of any interaction in transitive sentences is not surprising, given the null effects obtained in native speakers, and it can be explained away by the multifactorial nature of word order preferences (see Section 3.3 in Chapter 3 for a more detailed discussion of this issue). In ditransitive sentences, non-native bilinguals (L2Sp-L1Ba) present a *short-before-long* preference when ordering constituents of varying length, contrary to what they do in their native head-final language, Basque (Ros et al., 2015).

Crucially, the results in ditranstive sentences indicate that non-native speakers of Spanish behave differently from native speakers concerning the bias for DLM. Specifically, L2Sp-L1Ba bilinguals resort to shifted *short-before-long* S-V-IO-O orders less often than native speakers, even though they are extremely competent in their second language and were exposed to it from early in life. L2Sp-L1Ba speakers shift long constituents after short ones when using Spanish, unlike in their native language, Basque. However, they do so with a lower frequency than native speakers, either monolinguals or L1Sp-L2Ba bilinguals. The weaker proportion of shifted S-V-IO-O orders entails a higher proportion of S-V-O-IO canonical *short-before-long* word orders. This is a similar outcome to the one reported in Dennison (2008): When using Korean,

	Estimate	SE	z-value	p-value
Ditransitive sentences (N = 1427)				
Intercept	-2.218	.370	-6.003	<.0001 ***
Length: Long-IO	-.387	.363	-1.065	.287
Length: Long-O	1.318	.173	7.613	< .0001 ***
Group: Spanish monolinguals	.129	.462	.279	.781
Group: L1Spanish-L2Basque	.961	.430	2.236	.03
NP Position: IO-O	-.529	.209	-2.529	.01 *
Long-IO * Spanish monolinguals	-.032	.523	-.061	.951
Long-O * Spanish monolinguals	.884	.252	3.505	.0005 ***
Long-IO * L1Spanish-L2Basque	.711	.439	1.620	.105
Long-O * L1Spanish-L2Basque	.652	.243	2.680	.007 **
Long-IO * NP Position	.223	.296	.751	.453
Long-O * NP Position	.123	.120	1.021	.307
Spanish monolinguals * NP Position	-.259	.358	-.723	.469
L1Spanish-L2Basque * NP Position	-.427	.312	-1.369	.171
Long-IO * Spanish monolinguals * NP Position	.252	.510	.494	.621
Long-O * Spanish monolinguals * NP Position	-.041	.204	-.201	.841
Long-IO * L1Spanish-L2Basque * NP Position	.565	.426	1.328	.184
Long-O * L1Spanish-L2Basque * NP Position	-.230	.190	-1.208	.277

TABLE 4.10: **Logit mixed model analysis for ditransitive sentences in Spanish monolinguals and bilingual L1Sp-L2Ba and L2Sp-L1Ba bilingual speakers in Experiment 3**

balanced and L2Korean-L1English bilinguals speakers presented a higher proportion of S-O-V canonical *short-before-long* word orders than native speakers. Therefore, it is still not completely clear whether these results are due to a stronger preference for canonical orders as the result of the unbalanced trade-off between production facilitation and robust information transfer or to the transfer of *short-before-long* orders. In the following section, I turn to L2Ba-L1Sp bilinguals, whose behavior when using Basque can be useful in uncovering the reasons behind the weaker preference for *short-before-long* orders in L2Sp-L1Ba when using Spanish.

3 Experiment 4: Basque native and non-native speakers

My results in Experiment 3 clearly show that the preference for reducing dependency length varies across speakers. This confirms previous results from Dennison (2008). This is especially relevant, because the groups in Dennison (2008) were of different sizes and ANOVAs were used to analyze binary data. However, both Dennison's and my own study are inconclusive about the underlying cause of this variation. This issue is addressed in Experiment 4, by comparing the behavior of native and non-native speakers of Basque. Furthermore, this comparison can ultimately be informative of the inner-workings of the cross-linguistic tendency towards shortening dependencies. A crucial difference between Basque and Spanish with regard to word order preferences is that there is a greater variability regarding word order in Basque than in Spanish in all conditions, as shown in Experiments 1 to 3. In Basque, for instance, 42% and 33% orders were shifted in the All-Short and Long-IO conditions respectively (Ros et al., 2015, see Chapter 2). Therefore, there is more room to find any effect that involves comparisons between All-Short and Long-IO conditions in Basque than in Spanish. Next, I will briefly review the predictions that the the main hypotheses tested in this chapter make for non-native speakers of Basque.

If there is superficial transfer of the preferred length-triggered linear order from Spanish to Basque, then, in transitive sentences, L2Ba-L1Sp speakers should show no preference for DLM, since there is no effect in their native language. In any case, the effect of dependency length on Basque was quite small, reducing the possibilities of finding any differences in the expected direction (namely, an even smaller effect for L2Ba-L1Sp speakers). Crucially, in ditransitive sentences, I have more room in the Long-IO condition to confidently test the aforementioned hypotheses. Compared to L1Ba-L2Sp speakers, L2Ba-L1Sp speakers should tend to produce a higher number of *short-before-long* orders, namely shifted O-IO when the IO is long and canonical IO-O when the O is long. Crucially, in the Long-IO condition *long-before-short* and canonical orders do not overlap. With respect to verb position, L2Ba-L1Sp speakers will tend to produce verb-medial structures more often than L1Ba-L2Sp speakers, independently of constituent length, since that is the predominant linearization in their native language. Alternatively, if the magnitude of the tendency to minimize dependency length in a given language can carry over to a second one, in ditransitive sentences, L2Ba-L1Sp speakers

should produce a higher number of *long-before-short* orders in all conditions, that is, canonical IO-O when the IO is long and shifted O-IO when the O is long. This is so because the effect of DLM is greater in ditransitive sentences in Spanish than in Basque (see Chapter 3). That is, in ditransitive sentences, L2Ba-L1Spa speakers will present a stronger effect of DLM than L1Ba-L2Sp speakers. In transitive sentences, the opposite pattern is expected, a weaker or no effect of DLM in L2Ba-L1Sp speakers. With respect to verb position, L2Ba-L1Sp speakers will tend not to shift the verb in Basque, since they do not do it in their native language either, so that they will keep it in its canonical position at the end of the sentence. Last, if L2Ba-L1Sp speakers favor production ease over robust information transfer, they will produce more canonical orders than L1Ba-L2Spa speakers, especially when dependencies are long, which are assumed to be cognitively more costly. With regard to verb position, the prediction is similar. L2Ba-L1Sp bilinguals will tend to produce more canonical verb-final orders than L1Bas-L2Sp bilinguals, especially when dependencies are long.

3.1 Participants

28 high proficient L2Ba-L1Sp speakers were recruited and tested in Basque (20 females, mean age = 24, AoA = 2.04, SD = 1.31). 3 participants had to be discarded due to experimental error and 1 due to low competence in Basque. All participants were undergraduate students at the University of the Basque Country (UPV/EHU). They were all native speakers of Spanish and all had obtained a proficiency certificate in Basque equivalent to C1 level in the Common European Framework. Also, all their education previous to their undergraduate studies was conducted exclusively in Basque. Language background was tested using the same questionnaire as in Experiments 1 to 3 (see Section 1 in Appendix A). They rated their overall Basque proficiency level as 6.34 (SD = .55) and reported to use Spanish as their primary language for daily communications (4.62, SD = 1.49) (see Table 4.14). Self-assessed language use and proficiency for bilingual L1Ba-L2Sp participants in Experiment 1 (see Chapter 2) is reported in the Tables 4.11 and 4.13 below (see also Tables 2.1 and 2.2 in Section 2.2 in Chapter 2).

From the L2Ba-L1Sp participants, 3 came from an area with a density of bilinguals of 50-80%, 8 from an area with a 20-50% of bilingual speakers and 13 from an area where only up to 20% of the inhabitants are bilingual (IV Sociolinguistic Map, 2006). In the case of L1Ba-L2Sp participants, who participated in Experiment 1, 15 came from areas where bilinguals amounted to over 75% of the overall population, another 8 came from areas with 50-80% of bilinguals and 1 came from an area with less than 50% bilinguals (IV Sociolinguistic Map, 2006). Careful selection of the participants in this experiment is extremely important to ensure that non-native speakers of Basque were highly competent in both Spanish and Basque, due to the diglossic situation in the Basque Country. No monolingual speakers of Basque were included in this experiment because they are virtually non-existent in the Basque Country, especially in this

age range. All participants filled in a consent form in Basque and were naïve to the purpose of the experiment and received a monetary compensation for their participation.

	Basque	**Spanish**
Listening	6.83 (.38)	6.00 (.72)
Speaking	6.79 (.41)	5.00 (.88)
Reading	6.71 (.46)	5.67 (.86)
Writing	6.42 (.72)	5.17(.76)
Mean	6.69 (.49)	5.46 (.80)

TABLE 4.11: **Proficiency level scores (SD in brackets) as reported by L1Ba-L2Sp participants in Experiment 1**. *Self-assessed language proficiency is rated on a 7-point scale: 7 = Native-like level; 6 = High level; 5 = Medium-high level; 4 = Medium level; 3 = Medium-low level; 2 = Low level; 1 = No knowledge*

	Basque	**Spanish**
Listening	6.58 (.58)	6.88 (.44)
Speaking	6.00 (.93)	6.83 (.48)
Reading	6.58 (.58)	6.83 (.48)
Writing	6.21 (.66)	6.71 (.55)
Mean	6.34 (.55)	6.81 (.44)

TABLE 4.12: **Proficiency level scores (SD in brackets) as reported by L2Ba-L1Sp participants in Experiment 4**. *Self-assessed language proficiency is rated on a 7-point scale: 7 = Native-like level; 6 = High level; 5 = Medium-high level; 4 = Medium level; 3 = Medium-low level; 2 = Low level; 1 = No knowledge*

	Infancy (0-3 years)	**Primary school (4-12 years)**	**Secondary school (12-18 years)**	**Adulthood (after 18 years)**
Home	1.25 (.44)	1.17 (.48)	1.21 (.51)	1.21 (.51)
School		1.50 (.66)	1.71 (.55)	1.71 (.95)
Other		1.71 (.55)	1.83(.48)	2.04 (.46)
Equally comfortable in both languages: 1; *More comfortable with Basque:* 24				

TABLE 4.13: **Language history (SD in brackets) as reported by L1Ba-L2Sp participants in Experiment 1**. *Self-assessed language use is rated on a 7-point scale:* 1 = Only Basque; 2 = Mostly Basque, a few times Spanish; 3 = Mostly Basque but Spanish at least 25% of the time; 4 = Basque and Spanish equally; 5 = Mostly Spanish but Basque at least 25% of the time; 6 = Mostly Spanish, a few times Basque; 7 = Only Spanish

	Infancy (0-3 years)	**Primary school (4-12 years)**	**Secondary school (12-18 years)**	**Adulthood (after 18 years)**
Home	6.17 (1.17)	6.08 (1.44)	6.33 (1.01)	5.96 (1.49)
School		1.96 (1.04)	2.33 (1.05)	3.33 (1.76)
Other		5.29 (1.12)	4.89 (1.03)	4.58 (1.21)
Equally comfortable in both languages: 9; *More comfortable with Spanish:* 15				

TABLE 4.14: **Language history (SD in brackets) as reported by L2Ba-L1Sp participants in Experiment 4**. *Self-assessed language use is rated on a 7-point scale:* 1 = Only Basque; 2 = Mostly Basque, a few times Spanish; 3 = Mostly Basque but Spanish at least 25% of the time; 4 = Basque and Spanish equally; 5 = Mostly Spanish but Basque at least 25% of the time; 6 = Mostly Spanish, a few times Basque; 7 = Only Spanish

3.2 Materials and Procedure

Materials, task and stimulus presentation were similar to Experiments 1 to 3 (see Sections 2.3 and 2.4 in Chapter 2). The experiment was conducted exclusively in Basque.

3.3 Scoring

A total of 16 (3%) transitive and 45 (6%) ditransitive sentences were coded as miscellaneous and discarded from further analysis in the same conditions as in Experiments 1-3 (see Sections 2.5 and 2.5). As previously, changes in tense were permitted, as well as synonym substitutions and minor deletions. Scoring resulted in a total of 560 (97%) transitive and 675 (94%) ditransitive target sentences for L2Ba-L1Sp speakers. Note that in the case of L1Ba-L2Sp speakers, changes and additions scoring resulted in 563 (97.7%) target responses in the transitive sentences and 703 (96.5%) in the ditransitive sentences (see Section 2.5 in Chapter 2)

3.4 Analysis

As in previous experiments, the maximum random effect structure justified by the data was used, using model comparison between all the models that converged and did not contain correlations between the random effects equal to 1/-1 (Baayen et al., 2008). I report p-values estimated by the R-package lmerTest (Kuznetsova et al., 2017). All analyses were conducted using R (R Core Team, 2014) and the lme4 package (D. M. Bates et al., 2014). Transitive and ditransitive data were analyzed separately. For transitive sentences, first I fitted the data from L1Ba-L2Sp to a logit mixed model with Length as fixed effect, and with by-subject and by-item random intercepts, and with by-subject slopes for Length. Next, I fitted the data from both

groups of participants to a logit mixed model with Length and Group and their interaction as fixed effects, and with by-subject and by-item random intercepts and by-subject slopes for Length. For the transitive sentences, models with NP Position on the Screen as a predictor did not converge or were severely overffited. For ditransitive sentences, I fitted the data to two logit mixed models, a first one with only data from L2Ba-L1Sp speakers and a second one with data from both L1Ba-L2Sp and L2Ba-L1Sp speakers. In the first one, Length and NP Position on the Screen were included as a fixed effect and by-subject and by-item random intercepts were included. In the second one, Length, Group and NP Position on the Screen and their interaction were included as fixed effects, as well as by-subject and by-item random intercepts. In all these models, Length was Helmert-coded (the mean of the conditions were O and IO were long to when short) and NP Position on the Screen was centered. Additionally, a last logit mixed effects model was fitted to the data, with verb position as dependent variable and Length and Group and their interaction as fixed effects, with by-subject and by-item random intercepts and by-subject random slopes for Length. This model was run only for ditranstive and not for transitive sentences due to data scarcity (Jaeger, 2011) (see also Chapter 2).

3.5 Results

3.5.1 Transitive responses

In the set of transitive sentences, L2Ba-L1Sp bilingual speakers did not commit more errors when the O was long than when short ($p > .7$), but they did when the S was long compared to when short ($\hat{\beta} = 1.393$, $z = 2.078$, $p < .04$). As for the number of errors, L2Ba-L1Sp speakers did not differ from L1Ba-L2Sp speakers nor was there any interaction between group and condition (all ps > .8).

As shown in Table 4.15 and Figure 4.4, L2Ba-L1Sp speakers used predominantly canonical word orders. In all conditions, over 96.8% of all the orders produced by L2Ba-L1Sp speakers were canonical. In the case of L1Ba-L2Sp speakers, the overall proportion of canonical orders was 90%. This difference reached marginal significance in the model ($\hat{\beta}$ = -1.389, z = -1.852, p = .06). This overall difference was mainly triggered by the difference in the proportion of shifted orders produced by L1Ba-L2Sp and L2Ba-L1Sp speakers when the object was long. Whereas the rate of canonical orders was quite high for both groups in both the All-Short (L2Ba-L1Sp: 99.5%; L1Ba-L2Sp: 96.5%) and the Long-S condition (L2Ba-L1Sp: 96.7%; L1Ba-L2Sp: 94.5%), L2Ba-L1Sp speakers produced a higher number of canonical orders than L1Ba-L2Sp speakers in the Long-O condition (94.2% vs. 80.3%). In other words, compared to L1Ba-L2Sp speakers, L2Ba-L1Sp bilinguals produced a smaller number of shifted orders when the O was long than when short (20% vs. 6%). However, despite this numerical trend, no interaction was found between Length and Group ($p > .5$). which can be due to ceiling effects. That is, I find no difference across groups. Also, these results need to be interpreted cautiously because this model suffers from mild collinearity (rs > .6) (Jaeger, 2011).

Group	**Length condition**		
	All-Short	Long-S	Long-O
L1Basque-L2Spanish	7 (4%)	10 (6%)	37 (20%)
Miscellaneous	0	9	4
L2Basque-L1Spanish	1 (1%)	6 (3%)	8 (6%)
Miscellaneous	3	11	2

TABLE 4.15: **Raw data and percentages (in brackets) of shifted responses in transitive sentences produced by L1Ba-L2Sp (see also Table 2.5 in Chapter 2) and L2Ba-L1Sp bilingual speakers**

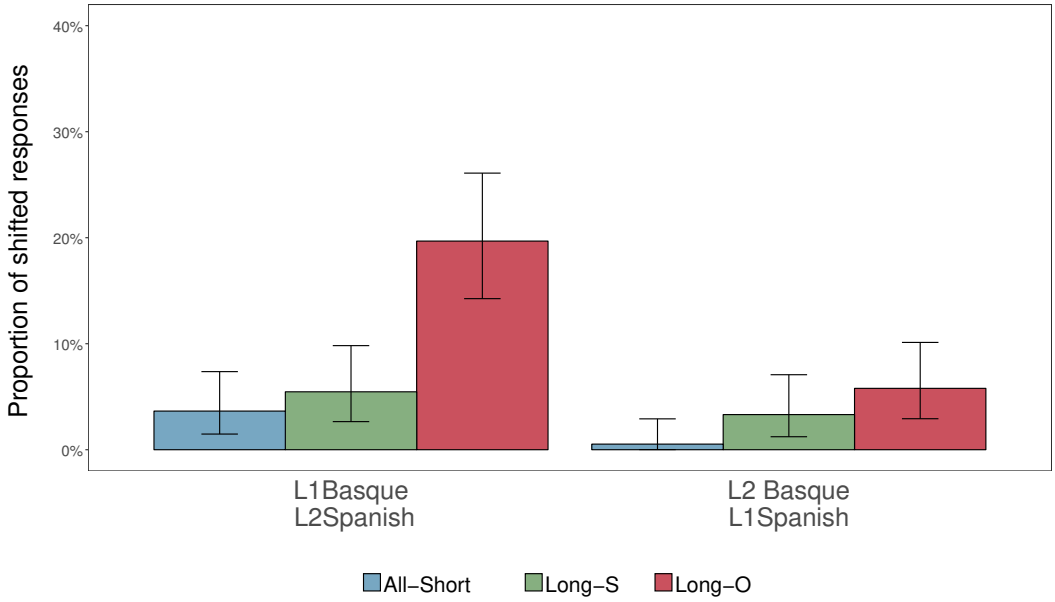


FIGURE 4.4: **Proportion of shifted responses in transitive sentences by L1Ba-L2Sp and L2Ba-L1Sp bilingual speakers in Experiment 4**. The error bars show 95% confidence intervals

	Estimate	**SE**	**z-value**	**p-value**
Transitive sentences (N = 560)				
Intercept	-5.182	1.347	-3.862	.0001
Length: Long-S	2.024	.1.912	1.059	.290
Length: Long-O	.756	.654	1.150	.250

TABLE 4.16: **Logit mixed model analysis for transitive sentences in L2Ba-L1Sp bilingual speakers in Experiment 4**

	Estimate	**SE**	**z-value**	**p-value**
Transitive sentences (*N* = 1123)				
Intercept	-3.861	.646	-5.976	<.0001 ***
Length: Long-S	.970	.711	1.365	.172
Length: Long-O	.533	.299	1.782	.08 .
Group: Non-native speakers	-1.389	.740	-1.852	.06 .
Non-natives * Long-S	.734	.756	.971	.332
Non-natives * Long-O	-.205	.315	-.652	.514

TABLE 4.17: **Logit mixed model analysis for transitive sentences in L1Ba-L2Sp and L2Ba-L1Sp bilingual speakers in Experiment 4**

3.5.2 Ditransitive responses

In the set of ditransitive sentences, L2Ba-L1Sp speakers did not produce more miscellaneous responses when the constituent was long than when it was short (all ps > .6). Also, the overall rate of miscellaneous responses was significantly higher for L2Ba-L1Sp than for L1Ba-L2Sp speakers ($\hat{\beta}$ = 1.560, z = 2.461, p = .01), but there was no interaction between native language and condition (all ps > .6).

As Figure 4.5 shows, L2Ba-L1Sp speakers placed the object before indirect object more often when it was long than when it was short, yielding a native-like *long-before-short* linearization preference ($\hat{\beta}$ = .664, z = 6.424, p < .0001, see Table 4.20 for the statistical analysis), in contrast to the *short-before-long* tendency shown in their native language, Spanish. When the IO was long, L2Ba-L1Sp bilinguals produced a 10% more canonical *long-before-short* IO-O orders than when all constituents were short, as was also the case for L1Ba-L1Sp speakers (Table 4.18). Also, L2Ba-L1Sp speakers were overall more prone to shift constituents when the object appeared to the left of the indirect object, in an O-IO reading order ($\hat{\beta}$ = .471, z = 2.293, p = .02, see Table 4.19 for the raw results and Table 4.20 for the statistical analysis), and this tendency did not interact with dependency length (all ps. > .2).

Group	**Length condition**		
	All-Short	**Long-IO**	**Long-O**
L1Basque-L2Spanish	99 (42%)	77 (33%)	140 (60%)
L2Basque-L1Spanish	71 (32%)	50 (22%)	121 (54%)

TABLE 4.18: **Raw data and percentages (in brackets) of shifted responses in ditransitive sentences produced by L1Ba-L2Sp (see also Table 2.7 in Chapter 3) and L2Ba-L1Sp bilingual speakers in Experiment 4**

Group	NP Position on the Screen	Length condition		
		All-Short	Long-IO	Long-O
L1Basque-L2Basque	*Inconsistent with shift direction*	32 (26%)	31 (26%)	62 (53%)
	Consistent with shift direction	67 (56%)	46 (40%)	78 (68%)
	Miscellaneous	4	4	6
L2Basque-L1Spanish	*Inconsistent with shift direction*	29 (25%)	22 (19%)	61 (54%)
	Consistent with shift direction	42 (39%)	28 (25%)	60 (53%)
	Miscellaneous	17	14	14

TABLE 4.19: **Raw data and percentages (in brackets) of shifted responses relative to position on the screen produced by L1Ba-L2Sp (see also Table 2.7) and L2Ba-L1Sp bilinguals in Experiment 4**

	Estimate	SE	z-value	p-value
Ditransitive sentences (*N* = 675)				
Intercept	-1.127	.367	-3.071	.002 **
Length: Long-IO	-.276	.191	-1.440	.150
Length: Long-O	.664	.103	6.424	.0001 ***
NP Position: O-IO	.471	.205	2.293	.02 *
Long-IO * NP Position: IO-O	-.209	.261	-.800	.423
Long-O NP Position: IO-O	-.172	.138	-1.241	.215

TABLE 4.20: **Logit mixed model analysis for ditransitive sentences in L2Ba-L1Sp bilingual speakers in Experiment 4**

	Estimate	SE	z-value	p-value
Ditransitive sentences (*N* = 1374)				
Intercept	-.915	.335	-2.733	.006 **
Length: Long-IO	-.003	.164	-.018	.985
Length: Long-O	.467	.090	5.207	<.0001 ***
Group: Non-native speakers	-.196	.477	-.411	.681
NP Position: O-IO	1.099	.185	5.946	<.0001 ***
Non-natives * Long-IO	-.269	.251	-1.072	.284
Non-natives * Long-O	.190	.136	1.395	.163
NP Position: O-IO * Long-IO	-.433	.224	-1.892	.06 .
NP Position: O-OP * Long-O	-.108	.128	-.845	.398
Non-natives * NP Position: O-IO	-.633	.275	-2.303	.02 *
Non-natives * NP Position * Long-IO	.216	.343	.630	.528
Non-natives * NP Position * Long-O	-.062	.188	-.328	.743

TABLE 4.21: **Logit mixed model analysis for ditransitive sentences in L1Ba-L2Sp and L2Ba-L1Sp bilingual speakers in Experiment 4**

Turning to the comparison between L1Ba-L2Sp and L2Ba-L1Sp speakers (see Tables 4.19 and 4.21 for the raw results and the statistical analysis respectively), all participants showed an overall tendency to reduce dependency length, by resorting to a *long-before-short* word orders when the O was long ($\hat{\beta}$ = .467, z = 5.207, p = < .0001). Also, both groups of speakers preferred to produce overall more shifted orders when they saw the IO constituent to the right of the O (O-IO) ($\hat{\beta}$ = 1.099, z = 5.946, p = < .0001). Crucially, native and non-native speakers did not differ from each other neither when the O nor when the IO were long ($\hat{\beta}$ = .190, z = .136, p = .163; $\hat{\beta}$ = -.269 , z = -1.072, p = .284., respectively). Also, there was an interaction between Group of speakers and Position on the Screen: L2Ba-L1Sp speakers produced less shifted O-IO orders than L1Ba-L2Sp speakers when O appeared in front of IO on the screen (in an O-IO order, a configuration consistent with the direction of shift in Basque) than when IO did (in an IO-O order) ($\hat{\beta}$ = -.633, z = -2.303, p = .02). Last, it is noteworthy that when NP Position on the Screen is not included in the model, the interaction between Group and Length reached marginal significance for the Long-O condition ($\hat{\beta}$ = .171, SE = .096, z = 1.789, p = .07), indicating a stronger tendency to shift long constituents before short ones for L2Ba-L1Sp speakers than for L1Ba-L2Sp speakers. This is probably due to the fact that, whereas the relative order of the constituents on the screen impacts all three conditions in the case of L1Ba-L2Sp speakers, in the case of L2Ba-L1Sp its effect is limited mostly to the All-Short condition (see Table 4.19). In L1Ba-L2Sp bilinguals, 30%, 14% and 15% more shifted O-IO orders are produced in the All-Short, Long-O and Long-IO condition, when the NP Position on the Screen is O-IO vs IO-O. In L2Ba-L1Sp speakers, this proportion reaches 14%, 6% and 1% in the All-Short, Long-O and Long-IO condition.

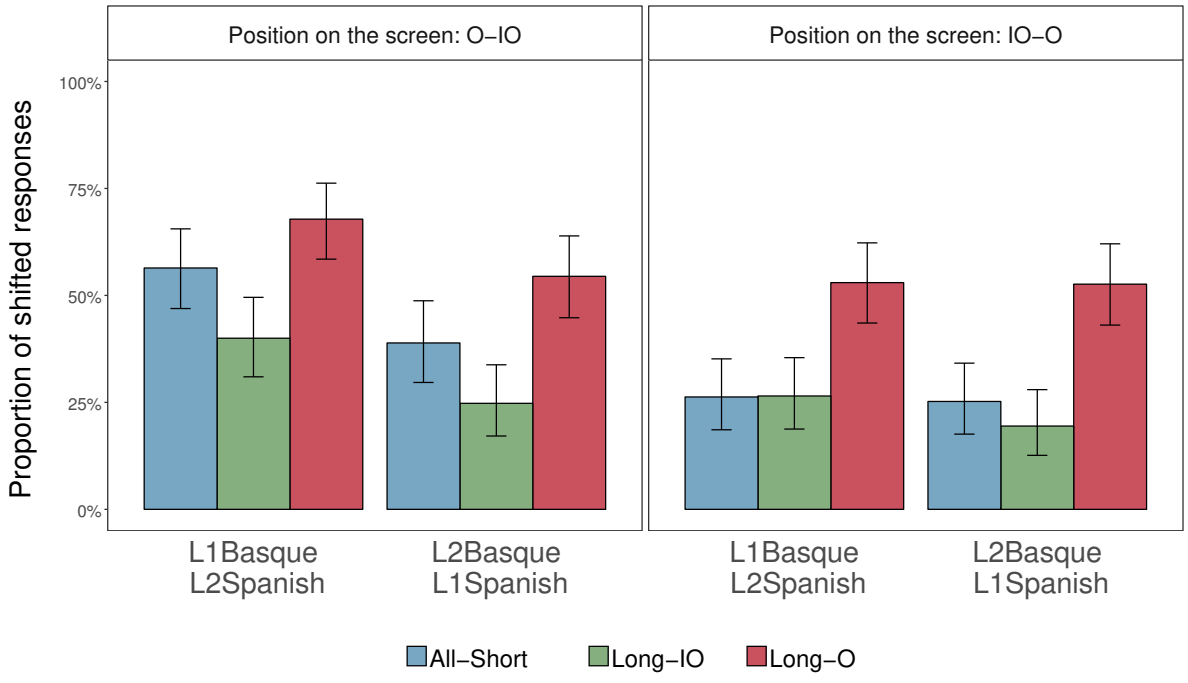


FIGURE 4.5: **Proportion of shifted responses in ditransitive sentences by L1Ba-L2Sp and L2Ba-L1Sp bilingual speakers in Experiment 4**. The error bars show 95% confidence intervals

To ensure that the effects of dependency length in non-native speakers were not due to a small subset of the items or the subjects, results by items and by subjects were examined. In the case of L2Ba-L1Sp bilinguals, of the 30 ditransitive sentences, 4 yielded no shift. Of the remaining 26, 22 yielded a number of shifts that placed the long phrase before the short one (see Figure 4.7). These results did not differ quantitatively from those of L1Ba-L1Sp speakers. In turn, of the total 24 L2Ba-L1Sp subjects, 4 did not show any preference to shift long elements relative to short ones. Of the remaining 20, 19 showed a tendency to shift long elements before short ones. Only 1 showed the opposite tendency to place long elements after short ones, less than in the case of L1Ba-L2Sp speakers (see Figure 3.5.2). As in Experiments 1-3, in both items and subjects there were numerical differences with respect to the magnitude of the effect of Length. I will address the differences across subjects in Chapter 5.

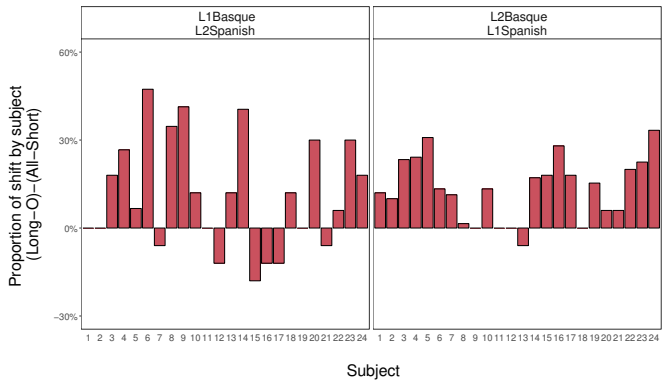


FIGURE 4.6: **Proportion of shifted responses by subjects in Experiment 4**. Negative proportions indicate a *short-before-long* ordering. No barplot indicates a preference for canonical orders. Positive proportions indicate a *long-before-short* ordering

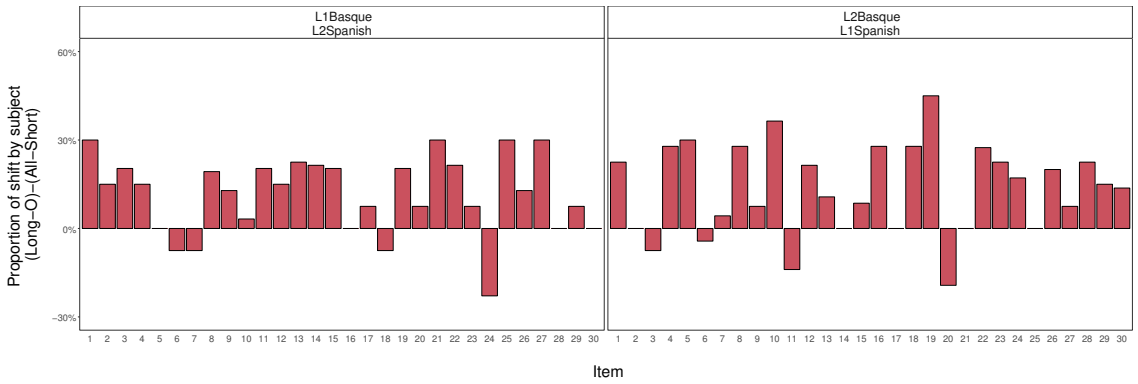


FIGURE 4.7: **Proportion of shifted responses by items in Experiment 4**. Negative proportions indicate a *short-before-long* ordering. No barplot indicates a preference for canonical orders. Positive proportions indicate a *long-before-short* ordering

3.5.3 Verb (and Subject) position

As stated in the Introduction, I was also interested in exploring L2Ba-L1Sp speakers' word order preferences with regard to the position of the verb in the sentence. Two logit mixed models were conducted to measure the number of verb-medial orders as a function of Length, in native and non-native bilingual speakers. Length was used as a predictor due to lack of enough data points to directly investigate the initial research question, that is, the impact of shifted non-canonical O-IO word orders on verb position (see Section 1). The raw data and percentage of verb-medial orders used when constituent order was non-canonical (O-S in transitive and O-IO in ditransitive sentences) are presented in Tables 4.22 and 4.24. For the sake of completeness, the raw data and percetanges relative to the use of verb-medial structures in the set of transitive sentences are reported in Table 4.22. However, due to data sparsity, no statistical analysis was conducted in this case. In Table 4.23 and Figure 4.8, verb-medial structures are shown as a function of constituent length in ditransitive sentences. This is the data fitted to the two models, whose results are reported in Tables 4.25 and 4.26.

Group	**Structure order**	**Length condition**		
		Al-Short	Long-S	Long-O
L1Basque-L2Spanish	O-S-V	2 (1%)	4 (2%)	9 (5%)
	O-V-S	5 (3%)	6 (3%)	28 (15%)
L2Basque-L1Spanish	O-S-V	0 (0%)	5 (3%)	3 (2%)
	O-V-S	1 (.5%)	1 (.6%)	8 (4%)

TABLE 4.22: **Raw data and percentages (in brackets) of shifted O-S responses relative to verb position in transitive sentences produced by L1Ba-L2Sp (see also Table 2.5 in Chapter 2) and L2Ba-L1Sp bilingual speakers in Experiment 4**

Group	**Length condition**		
	All-Short	Long-IO	Long-O
L1Basque-L2Spanish	97 (41%)	108 (47%)	134 (58%)
L2Basque-L1Spanish	78 (35%)	61 (27%)	89 (39%)

TABLE 4.23: **Raw data and percentages (in brackets) of verb-medial responses produced by L1Ba-L2Sp (see also Table 2.7 in Chapter 2) and L2Ba-L1Sp bilingual speakers in Experiment 4**

Group	**Type of response**		**Length condition**		
	Verb position	**Structure order**	Al-Short	Long-IO	Long-O
L1Basque-L2Spanish	Final	Others	7 (3%)	3 (1%)	11 (5%)
	Medial	S-O-V-IO	88 (37%)	71 (30%)	89 (38%)
		O-V-S-IO	4 (2%)	0 (0%)	36 (16%)
		Others	0 (0%)	3 (1%)	4 (2%)
	Total Final		7 (3%)	3 (1%)	11 (5%)
	Total Medial		92 (39%)	74 (32%)	129 (55%)
L2Basque-L1Spanish	Final	S-O-IO-V	3 (1%)	8 (4%)	34 (15%)
	Medial	S-O-V-IO	67 (30%)	42 (19%)	80 (35%)
		O-V-S-IO	1 (.5%)	0 (0%)	7 (3%)
	Total Final		3 (1%)	8 (4%)	34 (15%)
	Total Medial		68 (31%)	42 (19%)	87 (40%)

TABLE 4.24: **Raw data and percentages (in brackets) of shifted responses relative to verb position in ditransitive sentences produced by L1Ba-L2Sp (see also Table 2.7 in Chapter 2) and L2Ba-L1Sp bilingual speakers in Experiment 4** (verb-medial and verb-final orders with less than 5% occurrences in all conditions were subsumed under Others)

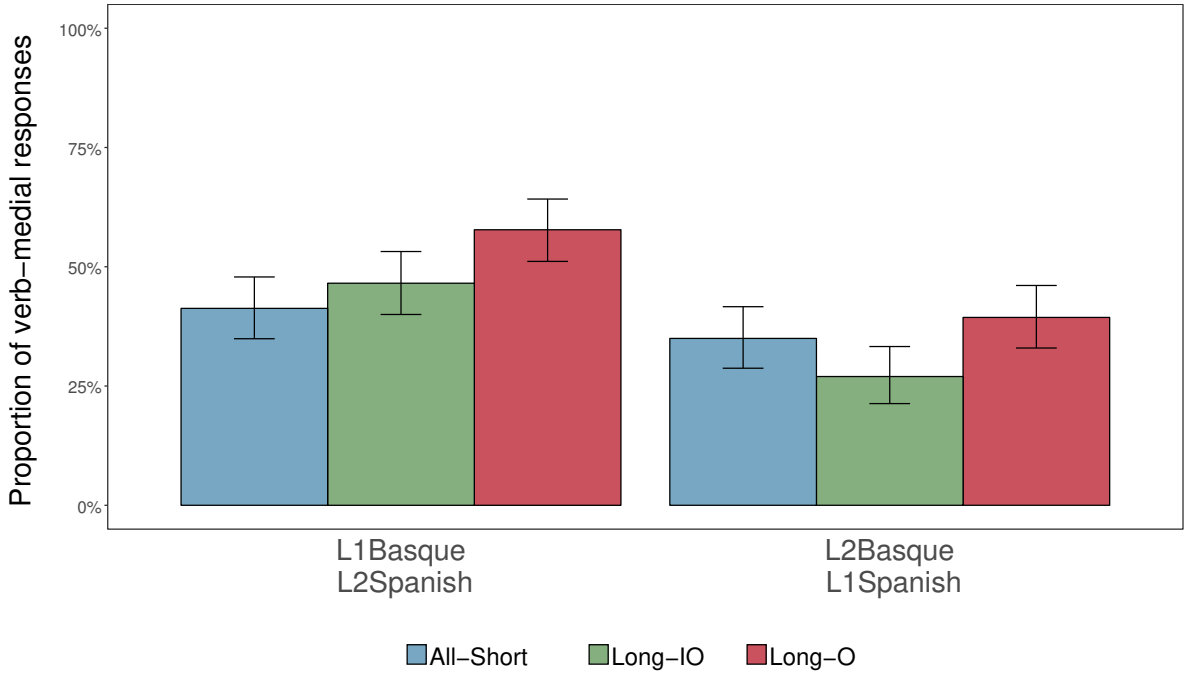


FIGURE 4.8: **Proportion of verb-medial responses in ditransitive sentences by L1Ba-L2Sp and L2Ba-L1Sp bilingual speakers in Experiment 4**. The error bars show 95% confidence intervals

Regarding the position of the verb in ditransitive sentences, as predicted, native speakers

use verb-medial structures more often when the O was long than when short (when a higher proportion of O-IO orders are produced), but not when the IO was long (where there is no difference with the All-Short condition with regard to the number of shifted orders produced) (see Table 4.23 and Figure 4.8 above and Section 2.7.3 in Chapter 2). This preference was interpreted as a strategy to speed up sentence processing and enhance communicative efficiency (see Chapter 2). L2Ba-L1Sp speakers do not show this tendency. They produce verb-medial orders only 4% more often when the O is long than when all constituents are short ($\hat{\beta}$ = .339, z = 1.303, p = .193). In contrast, they produced a greater number of verb-final orders when the IO was long than when short ($\hat{\beta}$ = -.728, z = -2.658, p = .007). In turn, when compared to L1Ba-L2Sp speakers, L2Ba-L1S bilinguals produced significantly less verb-medial structures when the IO was long, showing a stronger preference for canonical verb-final orders in this case.

	Estimate	**SE**	**z-value**	**p-value**
Verb position (*N* = 675)				
Intercept	-1.127	.550	-2.051	.04 *
Length: Long-IO	-.728	.274	-2.658	.007**
Length: Long-O	.339	.260	1.303	.193

TABLE 4.25: **Logit mixed model analysis for verb position in L2Ba-L1Sp bilingual speakers in Experiment 4**

	Estimate	**SE**	**z-value**	**p-value**
Verb position (*N* = 1374)				
Intercept	-.602	.488	-1.234	.217
Length: Long-IO	- .320	.300	1.069	.285
Length: Long-O	.999	.329	3.040	<.002 **
Group: Non-native speakers	-.478	.699	-0.684	.494
Non-natives * Long-IO	-.953	.454	-2.099	<.04 **
Non-natives * Long-O	-.756	.489	-1.546	.122

TABLE 4.26: **Logit mixed model analysis for verb position in L1Ba-L2Sp and L2Ba-L1Sp bilingual speakers in Experiment 4**

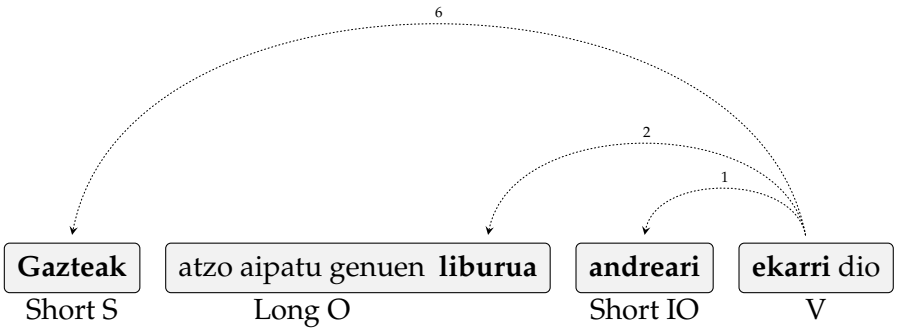
Furthermore, in ditransitive sentences, when the O is long, L2Ba-L1Sp speakers tended to produce shifted verb-final S-O-IO-V orders, where the subject appeared at the beginning of the sentence, 14% more often than when it was short. This structure was unattested in the production of L1Ba-L2Sp speakers, who, when resorted to shifted orders, produced O-V-S-IO orders also 14% more often when the O was long than when short (see Table 4.24) (note that the proportion of shifted orders produced by either group only reached 18%). In other words, when shifted orders were produced when the O was long, L2Ba-L1Sp speakers placed

the subject at the start of the sentence 94% of the time (n = 114 out of 121) of the time, while L1Ba-L2Sp did so only 71% of the time (n = 99 out of 140), as illustrated in Figure 4.9, where subject position is included for all types of orders.

An exploratory analysis was conducted to investigate whether, when dependencies were long and constituents were shifted relative to each other, native and non-native speakers produced fewer subject-first sentences and whether they differed from each other. Again, due to data scarcity, word order instead of length (since shifted orders have been shown to occur more frequently when dependencies are long), group and their interaction were used as fixed effects in the logit model. By-subject and by-item intercepts were used. The model with the interaction was marginally significantly better than without it (Df diff = 2, χ^2 = 4.9, p = .08). Post-hoc analysis of L1Ba-L2Sp bilinguals, on the one hand, and L2Ba-L1Sp speakers revealed that, while the first were affected by a preference to shift the subject to later positions in the sentences when the O was long ($\hat{\beta}$ = .688, z = .2.218, p = .03), L2Ba-L1Sp did not show such a preference ($\hat{\beta}$ = -.392, z = -.533, p = .594). This result, although post-hoc and exploratory, is quite relevant, since the shifted S-O-IO-V orders produced by L2Ba-L1Sp speakers, where the subject appears as the first element in the linearization and the verb as the last one, are less efficient than the shifted O-V-S-IO orders produced by L1Ba-L2Sp speakers, as shown in the examples 16a and 16b below, where sentences are represented as dependency graphs, to visually illustrate differences in dependency lengths across possible variants. Arches represent dependencies between the verbs and their arguments, measured in number of words. The sum of dependency lengths in the S-O-IO-V order amounts to 9, while it reaches only 5 in the O-V-S-IO order. This means that, when non-native speakers of Basque shift orders, they do it in a less efficient way than native speakers.

(16) **Shifted word orders produced by L2Basque-L1Spanish (16a) and L1Basque-L2Spanish (16b) bilinguals when O is long**

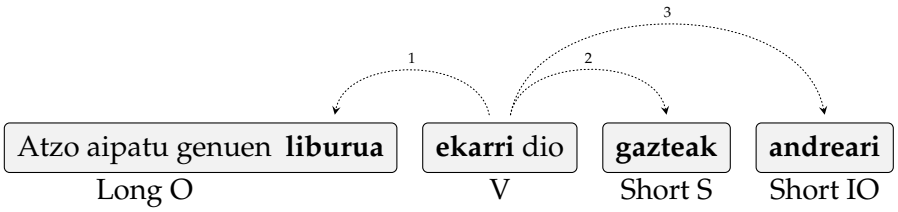
a. *Shifted order: S-O-IO-V*



Gazteak	atzo aipatu genuen liburua	andreari	ekarri dio
Youngster-the	yesterday mentioned-that book-the	lady-to	brought

The youngster has brought the lady the book that (we) mentioned yesterday

b. *Shifted order: O-V-S-IO*



Atzo aipatu genuen liburua	ekarri dio	gazteak	andreari
Yesterday mentioned-that book-the	brought	youngster-the	lady-to

The youngster has brought the lady the book that (we) mentioned yesterday

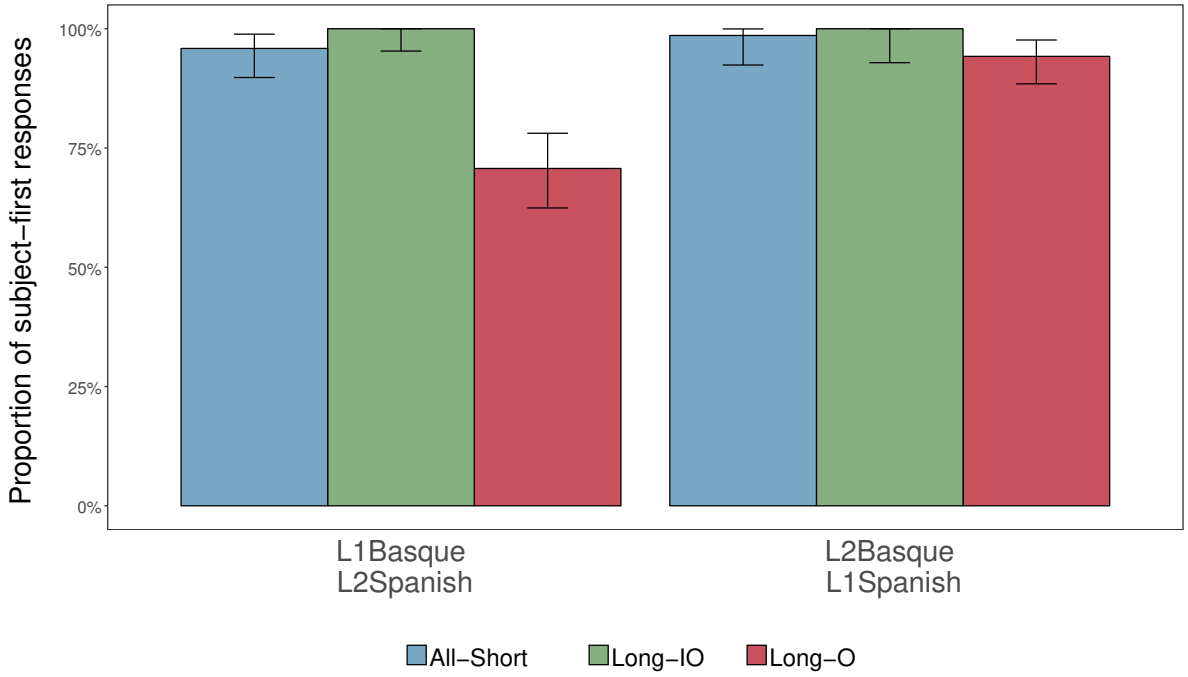


FIGURE 4.9: **Proportion of subject-first shifted responses in ditransitive sentences by L1Ba-L2Sp and L2Ba-L1Sp bilingual speakers in Experiment 4.** The error bars show 95% confidence intervals

3.6 Discussion

The results from Experiment 4 provide us with relevant evidence in delimiting the nature of the influence of the native language on length-triggered word order preferences in the non-native one. In transitive sentences, highly proficient non-native (L2Ba-L1Sp) bilinguals used predominantly canonical SOV orders in all cases and did not show any preference to reduce dependency length, but they did not differ significantly from native (L1Ba-L2Sp) speakers. In ditransitive sentences in Basque, L2Ba-L1Sp speakers employ a native-like *long-before-short*

strategy even though their native language, Spanish, shows a very strong opposite *short-before-long* preference. When the O is long, L2Ba-L1Sp speakers tend to produce shifted non-canonical O-IO orders. Furthermore, when the IO is long, they do not resort to *short-before-long* O-IO orders, as would be predicted if superficial word order transfer were the underlying cause of shifting differences between native and non-native speakers. Note that in this case, when IO is long, canonical and *short-before-long* orders do not overlap. Also, non-natives did not resort to DLM more often than native speakers, as predicted if abstract word order transfer were the cause of shifting differences across the two groups of bilinguals. With respect to verb position, L2Ba-L1Sp speakers produced more verb-final structures when the IO was long, yielding a higher number of canonical S-IO-O-V orders in this condition. Also, when they shifted orders, they produced a higher number of subject-first orders than their native L1Ba-L2Sp counterparts, yielding less efficient linearizations with longer dependencies.

My objective in this experiment was to collect more evidence on the possible causes of the differences in length-triggered word order preferences across native and non-native speakers reported in Dennison (2008) and in Experiment 3. Dennison (2008) argued that differences across groups with respect to the preference for short dependencies were due to superficial word order transfer from the preferred linearization present in the bilinguals' dominant language, which would constitute strong evidence against the generality of DLM as processing strategy. Some of the challenges faced by this hypothesis were outlined previously in this chapter. Crucially, the lack of a preference for a *long-before-short* order in transitive sentences in non-native bilingual speakers of Basque (Experiment 4) and Korean (Dennison, 2008) could be due to other causes: (i) abstract transfer of the null reliance on DLM in transitive sentences from their respective dominant language, English or Spanish, to the less dominant one, Korean or Basque (see Section 4.1 below); (ii) a preference for easier to produce canonical orders (see Section 4.2 below) when O is long, when *short-before-long* and canonical SOV orders overlap. The results obtained in ditransitive sentences in Experiment 3 indicated that non-native speakers were able to use the linearization preferred in their non-native language (in this case, *short-before-long* in Spanish), but they did so less frequently than non-native speakers. Crucially, the results obtained in ditransitive sentences in Experiment 4 confirm that the claim that non-native bilinguals transfer to their non-native language the linearization preferred in their respective native language is not accurate. Critically, in the conditions where a preference for the non-native Spanish-like *short-before-long* strategy does not coincide with a canonical linearization in Basque, that is, in Long-IO (where O-IO should be expected), I was not able to find any evidence supporting the hypothesis of superficial transfer. This strongly suggests that either the presence of abstract transfer or the preference for canonical orders are more feasible explanations for the pattern of results obtained across these three studies. Moreover, evidence from non-natives' word order preferences with regard to verb and subject position also point to reliance on canonical orders as the strongest candidate of these two hypotheses. In the next section, I discuss whether the pattern of results obtained in Experiments 3 and 4 together can be accounted for by any of the hypotheses outlined in the Introduction.

4 General discussion

In Experiments 3 and 4, I investigated whether bilingual non-native speakers differ from native speakers with respect to word order preferences as a function of dependency length. My overarching aim was to shed light on the inner workings and mechanisms behind the cross-linguistic bias for DLM. The results from these two experiments are compelling evidence for the universality of the principle of DLM. In Basque and Spanish, the shortest distance between dependents of varying length is obtained through opposing linearizations: *long-before-short* in Basque and *short-before-long* in Spanish. In Experiments 3 and 4, non-native bilingual speakers do not resort to the linearization favored in their respective native language when speaking their non-native one. That is, non-native speakers order constituents relative to their length exactly as native speakers.

However, Experiments 3 and 4 also demonstrate that the strength of the bias for short dependencies varies across speakers. When speaking their non-native language, even high proficient bilinguals quantitatively differ from native speakers with regard to word order preferences. As shown in Experiment 3, in Spanish, L2Sp-L1Ba speakers were less affected by the tendency to reduce dependency length than L1Sp-L2Ba speakers. In Experiment 4, in Basque, L2Ba-L1Sp bilinguals showed a stronger tendency for canonical verb-final orders than L1Ba-L2Sp speakers, especially when the IO was long, and for subject-initial orders, when shifted orders were used. In both cases, this led to non-native speakers employing word orders with longer dependencies. These results are in line with the idea that non-native speakers are less apt in maintaining a balanced trade-off between production ease and communicative success than native speakers. In the following sections, I will address in more detail the global pattern of results obtained in Experiments 3 and 4.

Aside from its theoretical relevance, whether there are differences between native and non-native speakers with regard to the preference towards DLM is a relevant issue for practical reasons, especially in a bilingual society as ours in the Basque Country. DLM is proposed to be an efficiency-triggered strategy. Short dependencies are claimed to be more easily processed. Across a wide number of languages, speakers have been shown to produce them with a proportion higher than chance. This has led to the hypothesis that language is efficiently designed for communicative purposes. In this regard, I was interested in whether non-natives' word order preferences would render a language more or less efficient than in the case of native speakers. Previous descriptive and prescriptive work in Basque has acknowledged the idea that there are differences across structures with respect to their degree of parsing complexity, and that L2Ba-L1Spa bilingual speakers are more prone to make use of the structures that are cognitively more costly (see Section 4.3). In Section 4.2 and 4.3, I review the empirical evidence for this claim based on the research covered in this chapter, where non-native speakers of both Spanish and Basque were compared with their native counterparts.

4.1 Differences across bilingual speakers as the result of abstract transfer?

When speaking Basque, high proficient early non-native speakers employ a native-like *long-before-short* strategy, even though their L1, Spanish, shows the opposite *short-before-long* preference. Similarly, when speaking Spanish, high proficient early non-native speakers resort to a *short-before-long* strategy, even though their L1, Basque, favors the opposite *long-before-short* tendency. This is unlike the results in Dennison (2008), where bilingual speakers of Korean whose dominant language was English or that were described as balanced did not shift long elements before short ones in transitive sentences. Also, in the case of non-native speakers of Spanish (L2Sp-L1Ba bilinguals) in Experiment 3, the tendency to produce *short-before-long* was weaker than in the case of native speakers of Spanish. I show, however, that these differences are not due to transfer of the dominant or native language's preferred superficial word order patterns to the non-native one, as argued by Dennison (2008). To arrive to this conclusions, data from Experiment 4 are crucial.

In Basque, L2Ba-L1Sp bilinguals do not show any tendency to produce Spanish-like *short-before-long* orders, neither when the O nor when the IO are long. The Long-IO condition in ditransitive sentences in Basque is the crucial test-case for the effect of dependency length on word order preferences in non-native speakers. In this condition, canonical IO-O and *short-before-long* O-IO orders do not overlap with each other. That is, when bilingual non-native speakers of Basque could have shifted constituents to comply with the *short-before-long* strategy favored in their native language, Spanish, without any overlapping with a canonical linearization, they do not, neither when O nor when IO is long. In other words, a seemingly *short-before-long* order triggered by superficial word order transfer only arises in non-native speakers when it overlaps with the canonical order in the language in question, as in Dennison (2008). The results from Experiments 3 and 4 reported in this chapter argue against the role of superficial transfer in determining length-triggered word order preferences in high proficient non-native speakers of head-final and head-initial languages.

Some pieces of evidence from Experiments 3 and 4 are suggestive of the role of abstract transfer in non-natives' word order preferences as the main factor behind my and previous results in the literature. In Spanish, L2Sp-L1Ba bilingual speakers showed a weaker preference towards short dependencies than native speakers. In Basque, although L2Ba-L1Sp speakers do not significantly differ from L1Ba-L2Sp speakers, they produced numerically more *long-before-short* orders (approximately 5% more). This is expected if the stronger preference towards DLM from Spanish overrides the weaker one found in Basque (see Chapter 3). In the same vein, L2Ba-L1Sp participants produced verb-final orders 14% more often than native speakers, as predicted if abstract transfer were a factor at play, although this difference was not significant. It is also noteworthy that the position of the O and the IO on the screen during the experimental task impacted speakers differently depending on their native language. In the case of L2Ba-L1Sp bilinguals, when the O is long, the proportion of shifting is not reduced when constituents appear on the screen in a canonical IO-O order, an order inconsistent with

the expected direction of shift in Basque (e.g., O-IO). In contrast, for native speakers, the proportion of shifted orders is equally reduced in all conditions when the canonical IO-O order is previously seen on the screen, also when O is long. This opens up the possibility that L2Ba-L1Sp speakers present a more robust effect of DLM than L1Ba-L2Sp speakers. Also, the rate of shifting across subjects seems more consistent across L2Ba-L1Sp speakers, although their exposure to Basque is likely subject to more variation than in the case of native speakers. Whereas only 14 L1Ba-L2Sp subjects (58%) resorted to *long-before-short* orders in Basque, 19 L2Ba-L1Sp (79%) did. However, although there is some supporting evidence for the claim that abstract transfer underlies non-native speakers' word order preferences as a function of length, more direct evidence should be desirable, since I failed to find any interaction between the two groups of bilinguals of Basque in Experiment 4. L2Ba-L1Sp speakers in this experiment do not present a stronger effect of DLM than L1Ba-L2Sp speakers.

There is one possible explanation for not finding any difference between native and non-native speakers of Basque with regard to the relative position of constituents of varying length. It could be argued that, compared to L2Sp-L1Ba speakers in Experiment 3, L2Ba-L1Sp speakers in Experiment 4 are more exposed to their non-native language. This means that any possible difference between native and non-native speakers of Basque with regards to the level of exposure to their non-native language is smaller than between native and non-native speakers of Spanish. In fact, L2Ba-L1Sp speakers in Experiment 4 reported a higher proficiency and more extensive exposure to their non-native language than L2Sp-L1Ba speakers in Experiment 3, as can be seen in the self-reported proficiency and exposure ratings provided in Tables 4.4 and 4.3 (Language history) and Tables 4.2 and 4.1 (Proficiency). This could be taken to mean that L2Ba-L1Sp speakers (Experiment 4) are more balanced bilinguals than L2Sp-L1Ba bilinguals (Experiment 3). Although this explanation would be fully compatible with the hypothesis of abstract transfer as the factor underlying any differences across native and non-native groups of speakers, I believe it to be incorrect. L2Sp-L1Ba speakers in Experiment 3 are very likely underestimating their abilities in, and exposure to, their L2, Spanish. Hakuta and d'Andrea (1992) present evidence that language attitude contaminates self-rated language proficiency scores. They claim that self-reported language ability is as much a measure of language attitude as it is of language proficiency. This might be relevant for my participants, since Basque and Spanish coexist in the Basque Country in a diglossic situation: There are no adult monolingual speakers of Basque, and only 32% of the population over 16 years in the Basque Country is highly proficient in both languages, whereas 50.8% are monolingual Spanish speakers (V Sociolinguistic Survey, 2011). Furthermore, recent surveys on language use in the school system indicate that L1Sp-L2Ba children and adolescents tend to use their non-native language in a much lesser extent than L1Ba-L2Sp children and adolescents (Department of Education and Culture, 2013b). Also, Spanish prevails in mass media. As a consequence, L2Sp-L1Ba bilinguals (Experiment 3) tend to be much more balanced than L2Ba-L1Sp bilinguals (Experiment 4), in spite of their

self-reported linguistic background and proficiency[2]. Further work where exposure to the non-dominant language is objectively measured is desirable to address this possibility.

4.2 Trade-off between production facilitation and efficient communication

Information and language processing by humans is claimed to be an efficient and rational system (Anderson, 1990; Jaeger, 2010; Piantadosi et al., 2012; Shannon, 1948; Zipf, 1949, *inter alia*). Speakers will try to reduce production effort in order to avoid redundancies in the signal that are not required for a correct interpretation of the intended message by the listener. DLM can be understood in these terms. Sentences with short dependencies are easier to process than those with long dependencies (Bartek et al., 2011; Gibson, 1998, 2000; Grodner and Gibson, 2005). At the same time, speakers of a wide number of languages have been shown to favor the production of short dependencies (Futrell et al., 2015b), which has been seen as evidence that DLM responds to efficient considerations (Fedzechkina et al., 2018; Futrell et al., 2015b; Hawkins, 2014). In the Introduction, I argued for the possibility that this trade-off between reducing production effort and ensuring robust information transmission might be unbalanced in non-native speakers, for which speaking their non-dominant language might be more costly than for native speakers.

The predictions made by this hypothesis are met in results reported in Dennison (Dennison, 2008) and in ditransitive sentences in Experiment 3, as well as (partially) in transitive sentences in Experiment 4. Crucially, in these experiments, non-native speakers produced a higher proportion of canonical orders either in one condition in particular (Dennison, 2008, Experiment 3) or overall (Experiment 4). This stronger preference for more frequent canonical orders found in speakers with less exposure to Korean, Spanish or Basque could be the result of the need to decrease production effort. This tendency should be especially visible in those cases where long dependencies and a higher production cost are involved. This explains the interactions in Dennison (2008) and in ditransitive sentences in Experiment 3. Moreover, this hypothesis is also compatible with non-natives' preference for canonical verb-final orders when the IO was long and for subject-first orders when shifted orders were used in ditransitive sentences in Experiment 4.

Additionally, under this account, it is not surprising that I fail to find any interaction between dependency length and native versus non-native speakers in both transitive and ditransitive sentences in Basque, while I do in Spanish. I might lack the statistical power necessary to find any effect and my results could be a Type II error. According to the

[2]Moreover, the rate of errors across groups in Experiments 3 and 4 are compatible with this claim. The number of errors produced was overall higher for L2Ba-L1Sp bilinguals speaking in Spanish in Experiment 3 than for L2Sp-L1Ba bilinguals speaking in Basque in Experiment 4 (see Tables 4.7 and 4.15), although no interaction between group and condition emerged. In any case, nothing indicates that L2Ba-L1Sp bilinguals faced fewer difficulties than L2Sp-L1Ba bilinguals when using their non-native language. However, I acknowledge that other measures of production difficulty were not collected (disfluencies, response latencies, etc.),

hypothesis discussed here, non-native speakers should show a weaker preference for shifted orders than native speakers, especially when dependencies are long. In Spanish ditransitive sentences, native speakers produced shifted orders approximately 75% more often when O was long versus when short. In Basque ditransitive sentences, native speakers shifted orders approximately 20% when O was long versus short (see Chapter 2). Any effect in the expected direction entails a smaller number of shifted orders to be produced by non-native compared to native speakers. It is quite likely that any differences in the production of shifted word order preferences is negligible given the small effect size of the preference for DLM in Basque and the sample size used in Experiment 4. This is also the case for the predicted higher proportion of verb-final orders in the condition when the O was long. Unlike in Long-IO, the interaction between verb-position and group would not be expected to emerge so easily, because in this condition shifted O-IO non-canonical orders tend to be produced with more frequency. Further research where enough statistical power is obtained is necessary in order to ascertain the likelihood of this second possibility.

It could be argued that the higher production of canonical orders found in Experiments 3 and 4 is due to a preference for syntactic simplicity (Erdocia, Laka, I., and Rodríguez-Fornells, 2009; Frazier, 1979; Kimball, 1973, among others). A number of behavioural and electrophysiological studies have shown higher processing costs in canonical versus non-canonical orders in languages that allow word order freedom, such as German, Japanese or Basque (see Erdocia et al., 2009 and references therein). Some of these studies have argued that this difference is due to the greater syntactic complexity of the non-canonical structures, also in production (Scontras et al., 2015). Derived orders involve more syntactic computations than their canonical counterparts and are claimed to be thus processed slower and with more cost (Frazier, 1990; Frazier and Fodor, 1978). Non-native speakers, who might find more costly to use their non-native language, might thus rely more often on canonical orders. However, all these studies face a clear confound. Canonical orders tend to be more frequent than their non-canonical counterparts. For example, in Basque, a language relevant to my research, OSV orders elicit longer reading times and a modulation of P600 component than SOV orders (Erdocia et al., 2009). At the same time, in a written corpus study from De Rijk (1969), SOV is the most frequent order, being used up to 57% of the time, whereas OVS and OSV amount to only 6% and 4% of the data. It is thus impossible to disentangle whether any effect found in this case is due to frequency or to syntactic simplicity. Recent research points to the first option as the most plausible one. Experience-based accounts predict processing difficulties as a consequence of how listeners' previous experience shapes language processing (MacDonald et al., 1994; Trueswell and Tanenhaus, 1993; Trueswell et al., 1993). In fact, the pattern of processing cost vary as a function of the frequency of a given structure across and within languages, even within the course of a single experiment. Reduced relative clauses cause processing difficulties in English (causing so-called *garden path* effects), whereas they do not in French (Frenck-Mestre, 2005). Relevantly, reduced relative clauses are much more frequent in French than in English. More crucially, even within the same language, namely English, repeated exposure to *garden path* sentences reduces their processing disadvantage with respect

to main clauses. Moreover, main clause canonical interpretations become cognitively more costly (as signaled by longer reading times) in environments where they are hardly encountered (Fine, Jaeger, Farmer, and Qian, 2013), even though they are syntactically more simple than the reduced relatives (for relevant literature, see also Fraundorf and Jaeger, 2016; Wells, Christiansen, Race, Acheson, and MacDonald, 2009 and references therein, among others).

4.3 Efficiency in non-native language use

One key question in the investigation reported in this chapter is whether non-native speakers use language in a less, equal or more efficient manner than native speakers. This is an issue that researchers interested in the acquisition of Basque as a second language, language instructors and style books writers have tackled (Kaltzakorta, 2007; Maia, 2015; Zubimendi and Esnal, 1993). It is argued that less expert speakers tend to increase dependency length in Basque, which yields higher processing load than their possible counterparts with shorter dependencies (Kaltzakorta, 2007; Maia, 2015; Zubimendi and Esnal, 1993). Among the possible ways by which dependency length is increased authors cite the tendency to place the verb in a final versus a medial position when long dependencies are involved, or to increase the distance between elements that are highly dependent with each other for a complete interpretation (see NPs in italics in example 17 below taken from Kaltzakorta, 2007, as cited by Maia, 2015). This phenomenon has been called *back-burden* (*atzerakarga*), and has been attributed to the influence of surrounding languages such as Spanish or French, or to prescriptive tendencies in the school system since the second half of the twentieth century (Altube, 1929). Among these prescriptive rules, Maia (2015) mentions specifically the recommendation that verbs need to appear at the end of the sentences, and subordinate sentences before their main verb. Also, a number of authors argue that the widespread use of the so-called *back-burden* in Basque is a communicative disadvantage with respect to surrounding languages such as Spanish or French (Amuriza, 2012; Aristegieta, 2009; Maia, 2015), which tend to reduce dependency length to a higher extent.

My research allows us to provide some insights on the phenomenon of *back-burden* or *atzerakarga*. First, long dependencies are produced by native speakers, and with a high frequency. Research on other verb-final languages shows that long dependencies can in fact lead to processing facilitation, mainly head-final ones (German, Hindi and Japanese) (Konieczny and Döring, 2003; Nakatani and Gibson, 2010; Vasishth and Drenhaus, 2011; Vasishth and Lewis, 2006). This means that, also in Basque, long dependencies might not pose such a problem as previously thought. This is still an open empirical question in the case of Basque. Second, placing long elements before short ones is, in fact, an efficient strategy, unlike what is implicitly stated from examples such as 17, where, as a possible solution to the cost caused by the long dependency, it is proposed that the long RC should be placed after the verb. Third, crucially, when native speakers of Basque shift long elements before short ones,

they also place the verb in a medial position and do not keep the subject at the beginning of the sentence, uttering an O-V-S-IO sequence. In contrast, non-native speakers in Experiment 4 tended to produced subject-first verb-medial S-O-IO-V sentences, a structure unattested in the data from native speakers. As argued in Section 3.5.3, this tendency leads to longer dependency lengths in the case of non-native speakers.

(17) *Donostiako alkateak* errauste plantaren behin-betiko kokagunearen erabakia
Donostia-LOC mayor-ERG incineration plant-GEN final location-GEN
ateratzeko bere taldearen arrazoiak aipatzen dituen *idatzi bat* eraman zuen
decision-ABS release-NOM his party-GEN reasons mentions AUX-REL document a
atzoko bilerara.
brought AUX yesterday-LOC meeting-ALL
The mayor of Donostia has brought a document that mentions the reasons his party has to release the decision with regard to the final location of the incineration plant

The research presented in this chapter can be seen as an attempt to start framing the questions we need to be able to answer so that we are in the position to fully address the issue of whether, until what extent and under which circumstances bilingual speakers use their non-native language less efficiently than native speakers.

5 Concluding remarks

Evidence from bilingual speakers of languages with opposing preferences in relation to the ordering of constituents of different length (*long-before-short* in Basque versus *short-before-long* in Spanish) provides us with valuable information about the nature of the cognitive bias for DLM. Contrary to previous claims (Dennison, 2008), high proficient non-native bilinguals do not carry over to their second language the either *long-before-short* nor *short-before-long* surface ordering present in their dominant language. Interestingly, even extremely proficient non-native bilinguals who have been exposed to their second language from very early on differ from monolingual and native bilingual speakers with respect to the degree to which they reduce dependency length. Non-native speakers of both Spanish and Basque produce long dependencies more often than native speakers. I argue that this outcome is best explained as the result of an unbalanced trade-off between production ease and robust information transmission by non-native speakers. Further controlled experiments where statistical power is assured are relevant to confirm this hypothesis and to further investigate its consequences from a listeners' oriented perspective. Also, a subset of my results can be accounted for by the effect of abstract transfer. In this respect, a relevant question open for future research is whether the pattern of results obtained across Experiments 3 and 4 will hold for less proficient non-native speakers, or whether the developmental trajectory that emerges when acquiring length-related word order preferences in a second language will entail a

different pattern than the one found in highly competent bilinguals. All in all, my research indicates that even high proficient bilinguals' language use is less efficient than native-language use, since it entails a more frequent use of long dependencies, and might demand an extra effort on the part of the listener. It is thus crucial to further investigate how tolerable long dependencies are across languages and how other strategies might be used to mitigate their cost.

6 Summary of findings

1. When using Spanish, L2Spanish-L1Basque speakers do not show any effect of DLM in transitive sentences. In ditransitive sentences, they present a *short-before-long* strategy in ditransitive sentences. That is, they behave as native speakers of Spanish, in spite of the opposite *long-before-short* preference present in Basque, their native and more extensively used language. Crucially, non-native bilinguals of Spanish whose first language is Basque present a *weaker* preference for reducing dependency length and thus a higher rate of long dependencies than their native counterparts.

2. When using Basque, L2Basque-L1Spanish show a *long-before-short* preference in ditransitive sentences. That is, they behave as native speakers of Basque, in spite of the opposite *short-before-long* strategy in Spanish, their native and predominantly used language. Nevertheless, compared to native speakers, non-native bilinguals of Basque produce more canonical subject-first and verb-final structures, which yield *longer dependency lengths* than their verb-medial sentences. Also, in transitive sentences, they produced almost exclusively canonical orders, unlike native speakers, who shift long elements before short ones, although the difference across groups did not reach significance.

3. The pattern of results obtained across Experiments 3 and 4 speak against the role of superficial word order transfer in length-triggered word order preferences in non-native bilingual speakers of Basque and Spanish.

4. The pattern of results obtained in Experiments 3 and 4 indicate a less efficient non-native versus native language use in both the case of Spanish and Basque bilingual speakers (e. g., a more frequent use of long dependencies). I argue that these results best fit the hypothesis that non-native speakers tilt the balance between production facilitation and successful transmission of the intended message in favor of the former. Further work is needed to confirm this claim and to investigate its possible interaction with the effect of abstract transfer, which can also account for some of the results reported in this chapter.

Chapter 5

The effects of memory demands and individual differences on DLM

Abstract

Since the beginning of the field of psycholinguistics, memory limitations have provided a number of cornerstone assumptions in the developing of different theories of language use (see Tanenhaus, 2004 for a review). However, to this day, little evidence has been collected to support these assumptions, especially in the area of language production. With respect to the preference for dependency length minimization (hereafter, DLM), some researchers claim that it arises due to limited memory resources. Others assume that it ensures easier recovery of the intended message and enables efficient communication, and it is thus strongly determined by comprehension needs. In this chapter, I report a production experiment where I manipulated memory load together with measures of individual differences (IDs) in working memory capacity (WMC) to explore the predictions made by these two views. Results indicate that a concurrent memory load interacts with speakers' tendency to reduce dependency length, but only when the relevant DLM-abiding order has been previously read on the screen. Also, reliance on DLM strategies increases along WMC as measured by a Reading span task. On the other hand, performance in an Operation Span task fails to predict word order preferences. These results are in contrast with any theory that attributes DLM to memory demands imposed by long dependencies and align better with the idea that the bias towards short dependencies is the result of communicative efficiency-related considerations. Taken into account that Reading Span tasks measure linguistic ability as well as WM capacity, my results may be linked to comprehension mechanisms. This study, which constitutes a first step towards a better understanding of the underlying causes of the preference for minimizing dependency length, underscores the need to refine our assumptions on the role of memory demands in language production and to further investigate the connection between language comprehension and production processes.

RESEARCH QUESTIONS:

- Do memory demands affect the preference for short dependencies?
- Does the preference for short dependencies interact with individual differences in working memory capacity?

1 Introduction

Understanding and producing language requires accessing information maintained in memory. This requirement is especially relevant when resolving long dependencies (Lewis et al., 2006). Also, individuals differ systematically with regard to memory capacity, and these individual differences (hereafter, IDs) correlate with language processing, acquisition and language final attainment (for a review, see Kidd et al., 2018). Crucially, assumptions on how WMC and IDs inform language processing vary across researchers and as times passes by, without receiving the necessary experimental scrutiny (for a relevant discussion, see Tanenhaus and Trueswell (1995, 2004)). Traditionally, in contrast to psychometric approaches, cognitive psychology and linguistics have disregarded the effect of individual variation on language processing and language form, which have been studied as monolithic phenomenona, whose core representations and processes were only marginally affected by individual variation (Kidd et al., 2018). With relation to WM, in their historical overview of the field until the mid-nineties, Tanenhaus and Trueswell (1995) describe how the view of human memory as a limited capacity system led to quite opposite (and untested) assumptions regarding its purported effect on the architecture of the parsing system[1]. Currently, the situation is improving in the field of language processing, and hypothesized memory mechanisms supporting cognitive operations during sentence comprehension are being directly tested, as well as their relation with individual variation along different cognitive abilities (see Subsection 1.2). In contrast, in language production, research focused on directly probing the theoretical assumptions regarding the interplay between memory demands and IDs, on one hand, and language production processes and representations, on the other, is virtually nonexistent (see Subsection 1.3). The general aim of this study is to join efforts with the few studies investigating WM and language production in opening up a very much needed research area into almost uncharted territory. Specifically, I focus on the role of memory demands on the bias towards DLM. To do so, I conducted a dual-production task where I manipulated dependency length and memory load and two complex span task to assess participants' WMC.

[1]Tannenhaus' view almost a decade later had not change and is still relevant to us in 2018. I reproduce it here (Tanenhaus, 2004: 377): "The relationship between WM and language processing, which contributes to motivate much theorizing, illustrates some of the limits of our current theories. [...] If any explanation for why processing is incremental is offered, it is that continuous processing and provisional commitments are necessitated by the demands of limited working memory. However, [...] working memory constraints could, and were, just as plausibility used to argue for delaying syntactic and semantic commitments."

In Subsection 1.1, I begin by providing a very brief overview of the two families of memory models that underlay past and current assumptions in psycholinguistic research. Next, in Subsections 1.2 and 1.3, I sketch the key findings on the role of memory processes in language comprehension and production that are relevant to my purposes, namely the study of the interplay between memory demands and language mechanisms when producing long dependencies. Specifically, in Subsection 1.3.1, I focus on implicit and explicit assumptions regarding the role of memory with respect to the bias towards short dependencies in production data and the predictions relevant for Experiment 5 that stem from them.

1.1 Working memory models: a snapshot

Memory effects on processing have been related to views of WM as either a series of peripheral storage systems and a central executive (Baddeley, 1986; Baddeley and Hitch, 1974) or as the activated portion of long-term memory, the focus of attention, in which most highly activated items are maintained (see Jonides et al., 2008 for a review of both type of models). In multi-core models as Baddeley (1986), items are accessed through a serial-search mechanism, whose activation in the specific storage systems (e.g., the phonological loop or the visuospatial sketchpad) decays over time, and can be boosted through active rehearsal. In contrast, in unitary-core models of memory, retrieval of intended items is assumed to involve parallel access to memory representations based on bundles of features that fit with those of the items maintained in long-term memory. Similarity-based interference among retrieval or encoding cues is the main cause of forgetting in this type of models (see Nairne, 2002), which may be also affected by activation decay. Along with the theoretical shift towards views of memory capacity in terms of interference-prone cue-based retrieval mechanisms, IDs with regard to WM are mainly understood as variation in individuals' capacity to allocate attention or to adequately retrieve and maintain information in the focus of attention, rather than exclusively derived from workspace limitations. Also, in multi-core models, some authors assume a unique system for language, analogous to the central executive in Baddeley's model (e. g, Just and Carpenter, 1992), whereas others argue for different and separate memory subsystems specific for different subcomponents or tasks related to language (e. g., Caplan and Waters (1999) assume two separate WM systems, a specific buffer for unconscious representations used in on-line comprehension, and another one dedicated to controlled verbally mediated tasks). Retrieval-models of WM assume a general and unitary system of human memory that interacts with the specialized task of language comprehension (Lewis et al., 2006).

There is a wealth of evidence supporting retrieval-models of WM. Specifically, in language comprehension, these models account for a great bulk of the experimental results reported in the literature and have also been directly probed (Gordon et al., 2004; Lewis, 1996; Lewis et al., 2006; McElree, Foraker, and Dyer, 2003; Van Dyke and Lewis, 2003, *inter alia*). The results from the few production studies that have tapped memory resources and syntax are also compatible

with them (Hartsuiker and Barkhusen, 2006; Slevc, 2011). In this dissertation, I thus assume a cue-based retrieval human memory system.

1.2 Working memory and individual differences in language processing

WM has played a fundamental part in language comprehension theorizing since the beginnings of the field. Miller and Chomsky (1963) argued that the number of embedded clauses that could be successfully parsed was constrained by upper limits intrinsic to human memory. Initially, models that proposed serial parsing as the most appropriate architectural feature of the processing system implicitly or explicitly assumed it to be the result of memory constraints (Frazier, 1979; Kimball, 1973). Modularity of language use was also related to arguments of memory capacity (Fodor, 1983, but see Just and Carpenter, 1992). Current memory-based accounts of locality effects such as the Dependency Locality Theory[2] (Gibson, 2000) or the activation-based model of human memory (Vasishth and Lewis, 2006) also assume that memory is a limited resource, even if the mechanisms involved in it may radically differ from previous models (Lewis et al., 2006). In any of these frameworks, the preference for short dependencies in language comprehension is a clear prediction. Computationally demanding sentences impose more taxing demands on memory. This means that processing long dependencies should be more costly for individuals under taxing memory conditions or with lower WMC. In the following paragraphs, I will very briefly review the most relevant findings regarding this prediction. Therefore, only studies that directly test this claim are included. All studies for which this claim is central but remains untested are excluded from this section.

King and Just (1991) found that individuals with lower reading span scores (low-WMC individuals) showed longer reaction times than their high-span counterparts (high-WMC individuals) when dependency distance was longer (that is, in object versus subject RCs in English)[3]. In garden-path sentences, MacDonald, Just and Carpenter (1992) showed that high-WMC individuals presented longer reading times when the temporary syntactic

[2] Although in DLT the preference for local dependencies in comprehension is memory-based, their assumptions regarding the memory processes involved remain language-specific and are captured via the so-called *storage* and *integration costs*. Using a simple word-counting measure of memory-based complexity, the *integration cost* is the sum of the number of words that intervene between a given head and a previous dependent, as Hawkins (1994) and myself thorough this dissertation. The *storage cost*, in turn, is the number of referents (non-given and definite nouns and finite verbs) that need to be maintained in memory until a dependency is resolved. Gibson himself (Gibson, 2000: 120) acknowledges that the *integration cost* is the part of the DLT theory that has received most attention and empirical support and it can be ultimately understood as a highly simplified proxy that reflects the cognitive processes of the memory system that underlay language comprehension and production, such as the ones proposed by activation-based models of human memory (Lewis et al., 2006). As mentioned in the main text, other researchers have used other metrics for sentence complexity, aside from number of words or intervening items, such as number of embeddings (Miller and Chomsky, 1963) or incomplete dependencies (Gibson, 1998; Kimball, 1973, *inter alia*)

[3] However, according to Waters and Caplan (1996), they do not report the results of the statistical analyses that would be necessary to demonstrate an interaction between group and dependency distance.

ambiguity was resolved towards a simple main verb sentence, compared to an unambiguous control sentence, while low-WMC individuals show no difference in reaction times but a lower accuracy rate in the ambiguous versus the unambiguous condition (see also Pearlmutter and MacDonald, 1995; for a correlation between WMC and the strength of the garden-path effect, see Farmer, Fine, Misyak, and Christiansen, 2017). This result is consistent with an interpretation based on memory as resource-limited (Just and Carpenter, 1992, but see Pearlmutter and MacDonald, 1995): When sentences are ambiguous, unlike low-WMC individuals, high-WMC individuals can construct and maintain in memory several syntactic interpretations in parallel, quickly and successfully setting for the correct one.

Three relatively recent studies provide additional evidence for the relationship between WMC and the preference for short dependencies. In all three, unlike the ones reported above, other measures of memory span aside from the Reading span task were used, as well as a continuous-variable instead of a extreme-groups design (Nicenboim, Logačev, Gattei, and Vasishth, 2016; Nicenboim, Vasishth, Gattei, Sigman, and Kliegl, 2015; Swets, Desmet, Hambrik, and Ferrerira, 2007; see Conway et al., 2005 and Baayen, 2004 for the benefits of a continuous-variable over an extreme-groups design). In a study on the preferences for low- vs. high- attachment in English and Dutch, Swets et al. (2007) found that low-WMC participants showed a greater preference for the more distant attachment site, whereas high-WMC participants preferred the more recent one. This finding was linked to memory-based differences with respect to prosody-based chunking strategies. High-WMC readers could maintain larger segments of text in memory, leading to a low attachment preference. Consistent with this result, in a study investigating the processing of long dependencies in German and Spanish, in an eye-tracking experiment, high-WMC readers showed the strongest locality effects, which were reduced as WMC decreased (Nicenboim et al., 2016). In contrast, in a self-paced reading task, participants with lower WMC showed a speed-up, which is normally interpreted as a conductual signature of *anti-locality* effects. In a previous study in Spanish, Nicenboim et al. (2015), also conducted a self-paced reading task and an eye-tracking study to investigate the processing of long dependencies in Spanish. In the eye-tracking experiment, as predicted by memory-based accounts of locality effects and unlike in Nicenboim et al. (2016), locality effects decreased as WMC increased until they became increasing *anti-locality* effects. In the self-paced reading task, however, low-WMC readers, unlike high-WMC ones, presented a speed-up when faced with long dependencies. The authors assume different underlying cognitive processes in the speed-up found in high- and low-WMC readers in the eye-tracking and the self-paced reading respectively in the two studies (Nicenboim et al., 2016; Nicenboim et al., 2015): The speed-up in low-WMC participants is interpreted as consistent with good-enough parsing strategies (F. Ferreira, Bailey, and Ferraro, 2002), and as the signature of *anti-locality* effects in high-WMC readers.

Some studies on the relationship between non-local dependencies and memory demands consist in dual-tasks, in which an on-line sentence processing and a memory task are

performed simultaneously[4]. In Gordon, Hendrick, and Levine (2002), participants read subject versus object cleft-sentences, which contained either personal names (e.g, *Mary*) or professions (e. g., *teacher*), while asked to remember a set of three words that matched either the occupations-condition or the proper names-condition. Participants were less accurate in subsequent comprehension questions about the content of the sentences in the match versus the no-match condition in the object-extracted sentences, compared to the subject-extracted sentences. In a design similar to Gordon et al. (2002), Fedorenko, Gibson, and Rohde (2006) also manipulated the number of words to be recalled (1 vs. 3), which were either similar or dissimilar to the nouns presented in the subject and object relative sentences employed in their study. They found that similarity between the nouns in the sentences and in the memory load affected the reactions times in the object RC to a greater extent than in the less complex subject RC (see also Fedorenko, Gibson, and Rohde, 2007).

Summing up, in most of the studies reported above, a relationship between WMC and/or memory load and different measures of processing cost has been repeatedly reported, which, to a certain extent, support the predictions from accounts that regard the preference for local dependencies as the outcome of the architectural characteristics of the human memory system (Gibson, 1998, 2000; Lewis et al., 2006). When processing more complex sentences which involve long dependencies, memory can be taxed through overload of the available computational resources or through higher retrieval or encoding costs either due activation decay or enhanced interference (Vasishth and Lewis, 2006). Nevertheless, a number of results described in this section are less consistent with this hypothesis than initially expected (MacDonald, Just, and Carpenter, 1992; Nicenboim et al., 2016; Nicenboim et al., 2015; Swets et al., 2007), and certain assumptions need to be made to accommodate the unexpected patterns into current models of language processing. One possibility is that, as argued by Nicenboim et al. (2016), the assumption that processing difficulty is unproblematically translated into longer RTs and facilitation into shorter RTs might be too simplistic. Also, other factors might interact with locality preferences (and, thus, obscure their effect), such as the role of memory, if any, in expectation-based processing. It is plausible that high-WMC individuals may have a more precise expectation of upcoming linguistic material, or may be able to maintain the relevant predictions in WM for a longer time or to retrieve them less costly than their low-WMC counterparts (see Nicenboim et al., 2015, or Farmer et al., 2017).

1.3 Working memory and individual differences in language production

In contrast to the situation in language comprehension research, the role of WM in language production has been barely addressed. There is a number of studies on the relationship between memory and lexical access (Daneman, 1991; Daneman and Green, 1986), written

[4] A few studies report an interaction between memory load and syntactic complexity in *off-line* measures (Caplan and Waters, 1999; Wanner and Maratsos, 1978). I do not cover them in this section because I focus on studies testing the effect of concurrent memory load during *on-line* language processing.

language (Kellog, 2004), agreement processes (Hartsuiker and Barkhusen, 2006; Slevc and Martin, 2016) and the scope of language planning (A. E. Martin and Freedman, 2001; A. E. Martin, Miller, and Vu, 2004; Wagner, Jescheniak, and Schriefers, 2010). However, to the best of my knowledge, the only study that investigates the role of memory demands in word order preferences is Slevc (2011) and no study has ever explored the relation between WM and the preference for word orders that yield short dependencies in language production, even though memory has been claimed to be the main underlying constraint in this preference (Liu, 2008; Scontras et al., 2015, among others).

Currently, accessibility effects are key in all theories of word order in psycholinguistics and in consensus models of language production (see also Chapter 2). In availability-based models of language production, it has been claimed that words compete for workpsace in the memory system, and more readily accessible words are uttered first in order to reduce the need to hold already-retrieved information and start producing early, which would help fluency (Bock and Irwin, 1980; Bock and Warren, 1985; Branigan et al., 2008; V. S. Ferreira and Yoshita, 2003; V. S. Ferreira and Dell, 2000; Prat-Sala and Branigan, 2000; Tanaka et al., 2011; Yamashita and Chang, 2001, among others). Conceptual accessibility is the ease with which a mental representation of a given referent can be activated in memory or retrieved from it. It can be inherent due to intrinsic semantic characteristics such as concreteness (Bock and Warren, 1985), animacy (Bock and Cutting, 1992; V. S. Ferreira and Yoshita, 2003; MacDonald, 2013), etc. It can also be derived from earlier use (Prat-Sala and Branigan, 2000). However, as for today, the only evidence bearing on the nature of accessibility effects and how they affect grammatical encoding is Slevc (2011), whose results are not straightforward. In Slevc, 2011, in a production-elicited experiment, when participants were under a verbal WM load, accessibility effects were reduced. Although these results are at odds with a view of accessibility effects as the result of the need to release space in working memory, Slevc argues that these results are compatible with the idea that accessibility effects on word order are due to the need to quickly produce early retrieved information that could otherwise be subject to interference[5]. In any case, this study shows that word order is vulnerable to manipulations of memory demands.

1.3.1 Working memory and DLM

Slevc (2011) provides evidence that backs up the idea that during grammatical encoding in production, as in comprehension, most operations hinge on the need to retrieve information from recent memory. The information maintained in or retrieved to the focus of attention during language production planning processes is susceptible to interference and less active

[5]In this study, no information is provided about the nature of the words presented to the participants, aside from their being unrelated to previous materials. It is thus unclear how similar they were to one or the other relevant NP in the target sentences. If accessibility effects are due to similarity-based interference during encoding or retrieval, and participants were presented with words equally similar to either NP in the VP, the element with a higher base activation (the more accessible one) should still be retrieved faster, according to activation-based models.

elements in memory experience decay (Gennari et al., 2012; MacDonald, 2013, *inter alia*), as described in Section 1.1. The effects of interference and decay should be especially clear in the case of long dependencies, where elements need to be switched in and out of the focus of attention (see Fedorenko, Gibson, and Rohde, 2006 for similar claims for comprehension). In fact, some researchers assume that DLM responds to memory demands to reduce the number of elements to need to be remembered while resolving a dependency (Ferrer i Cancho, 2004; Futrell et al., 2015b; Liu, 2008). However, as I have argued in previous chapters of this dissertation, DLM can also be understood as the result of efficient communication (Fedzechkina et al., 2018; Futrell et al., 2015b; Gildea and Jaeger, 2015; Hawkins, 2014). Current evidence upholds either view equally and only indirectly, as I will review next.

In two elicited production experiments, Scontras et al. (2015) found that non-local dependencies in RCs and wh-questions present longer latencies and word durations, as well as higher rates of disfluencies, than local dependencies. This is consistent with previous research that suggests that syntactically less complex speech is produced when under cognitive load, especially in individuals with low-WMC (Norman, Kemper, Kynette, Cheung, and Anagnopoulos, 1991). Additionally, corpus and experimental studies have attested a cross-linguistic preference for short versus long dependencies (Futrell et al., 2015b; Hawkins, 1994; Liu, 2008), as reported in Chapter 3. Word durations, error rates and frequency when employing a given structure, if there are other alternative candidates available, are all usually considered plausible evidence for the claim that producers experience difficulty when using long dependencies (but see also MacDonald et al., 2016 and a reply in Scontras, Badecker, and Fedorenko, 2017). However, differences in duration, latency or frequency between two or more variants of the same structure could also be due to other factors. Structures that are less frequent could be less active in memory and harder to retrieve, which would lead to longer durations and latencies (MacDonald, 2013). Crucially, they could be less frequent not because they are cognitively more costly, but because they are less reliable when conveying a given meaning. Given that long dependencies are harder to comprehend (see Section 1.3.1), this idea seems plausible. In fact, modulation of word durations and of frequencies of syllables, words, case endings and other elements have also been profusely attested in studies appealing to the notion of efficient communication, that is, of the tendency to provide maximum information value with the least cognitive effort (Aylett and Turk, 2004; Fedzechkina et al., 2012; Ferrer i Cancho and Solé, 2003; Jaeger, 2010; Piantadosi et al., 2011; Zipf, 1949, *inter alia*).

In this study, I aimed at testing the assumption that long dependencies are a burden for memory in production. Crucially, the views of DLM as either the result of memory limitations or of requirements of efficient communication lead to opposite predictions regarding the role of WM in length-triggered word order preferences. If the preference for DLM arises as a result of memory limitations, a WM load will increase the production of orders that yield short dependencies and WMC will correlate negatively with the preference for DLM. On the contrary, under the assumption that DLM is due only to communicative pressures for robust information transmission, which boils down to a trade-off between production difficulty

when producing shifted orders and the probability of communicative success (which is higher when using shorter dependencies), and provided that language production is more costly when memory demands are high, a WM load will decrease the proportion of orders that yield short dependencies and WMC will correlate positively with the preference for DLM.

2 Overview of Experiment 5

In Experiment 5, I investigated the role of memory demands on length-triggered word order preferences. In particular, I was interested in whether the preference towards reducing dependency length by producing a higher number of shifted long-before-short orders when long constituents are involved was affected by limitations in WM. A dual-task production experiment was conducted, where native speakers of Basque, who have shown a preference for DLM-abiding *long-before-short* orders, participated (see Chapter 2 and Ros et al., 2015). I assessed the number of *long-before-short* shifted orders participants produced when under more and less stringent memory conditions. Additionally, two complex span tasks, a Reading and an Operation span task adapted from Unsworth et al. (2005), were administered. The aim of these two WM tasks was to investigate how the responses in the production experiment were modulated by participants' individual WMC. I employed ditransitive sentences, where the length of the relevant constituents was manipulated (short vs. long) by adding a RC to either the O or the IO. For the predictions relevant for both the dual-task production experiment and the span tasks, see Section 1.3.1. The experiment was divided in two sessions, each one conducted a different day. In the first one, participants completed the dual-task cued-recall production experiment (see Section 2.2). In the second one, the two complex span tasks were administered (see Section 2.3). In the following sections, I first describe these two sessions separately (see below 2.2 and 2.3), and then turn to the results and the discussion of the main findings.

2.1 Participants

57 native speakers of Basque (n =1 excluded because she reported Spanish as her native language and n = 1 excluded due to equipment malfunctioning), all of them undergraduates at the University of the Basque Country (UPV/EHU) participated in this study (34 females, mean age = 22.45). Participants signed an ethical consent form upon arrival to the lab and were compensated economically for their collaboration. In the cue-based recall production experiment, participants' mean accuracy was 79.2% (SD = 14.5). Participants whose accuracy was below 70% were omitted from any analysis (n = 15). 3 more had to be removed because they committed more than 12 errors in the Operation span task and 1 because s/he did not complete all the experimental sessions. One last participant had to be removed from further analysis because s/he reported that s/he was under medication due a condition related to

memory after the experiment was conducted. This left us with a total of 37 subjects. Participants all signed an ethical consent form and they were compensated economically for their participation.

2.2 1. session: Cue-based recall production task

A cued-recall production experiment was conducted, based in the design used in Experiments 1 through 4. This first session lasted approximately 35 to 50 minutes.

2.2.1 Materials

32 ditransitive sentences (adapted from the materials used in Experiment 1) and 60 fillers were created (see Table 5.1 and Appendix 2). Four conditions were created by manipulating the Length of the object of the sentences to be recalled (All-Short vs Long-O) and the concurrent Memory load (Low vs. High). As in all experiments before, Length was manipulated by attaching a RC to the relevant constituent. All RCs in this experiment had a total of 6 words. Unlike in the experiments reported in previous chapters, sentence fragments were followed by auditorily presented words, which had to be recalled immediately after producing the target sentences. Either 1 or 3 words were presented to the participants in the Low and in the High Memory load conditions, respectively. In the Low Memory load condition, only the first word of each of the 32 sets listed in Appendix 3 was presented to the participants. Of the total of 96 words presented to the participants, 84 were animate and 12 inanimate. All were high frequency words (here average number of appearances in corpus) and had a length of 3-4 syllables (the determiner - *a* - is used in the lexical citation form in Basque, so that words with less than 3 syllables are uncommon in this case). They were read by a native speaker of Basque and edited using Audacity (Audacity Team, 2016). Last, in order to avoid effects due to reading habits, the position of the constituents on the screen was counterbalanced. In half of the items, the O appeared to the left of the IO, in a O-IO reading order. In the other half, the order of the items was reversed (IO-O).

2.2.2 Procedure

Four lists were constructed and each participant was presented with one list, in an individually pseudorandomized order, with the constraint that no more than two experimental sentences were presented next to each other. The experiment was performed using Presentation software (version 16.3) (Neurobehavioral Systems Inc., 2016). The participant pressed the spacebar to start a trial. This triggered the display of the first screen with four boxes containing the unordered sentence constituents. The verb appeared always in

a box at the left-upper corner of the screen. Participants had up to 8 seconds to read the unordered constituents and to prepare a sentence. If they were ready faster, participants could press the spacebar to continue to the next screen. Next, they heard either 1 or 3 words before the verb was automatically presented in the middle of the screen and they had to produce the sentence previously seen and, immediately after that, the word(s) (see Figure 5.1). Unlike Experiments 1-4, Experiment 5 was not presented under the guise of a recognition task and, in consequence, no fillers were repeated.

Phrase length condition	**Example**
All-Short	Amak irakasleari pastela eraman dio Mother$_{Erg.3sg}$ teacher$_{Dat.3sg}$ cake$_{Abs.3sg}$ take AUX The mother brought the cake to the teacher
Long-O	Amak irakasleari [arratsalde osoan etxeko frigorifikoan izan dugun pastela] eraman dio Mother$_{Erg.3sg}$ teacher$_{Dat.3sg}$ [afternoon whole-the house-of fridge have AUX-that cake$_{Abs.3sg}$] take AUX The mother brought the cake that we kept in the fridge at home the whole afternoon to the teacher

TABLE 5.1: **Example set of experimental ditransitive sentences used in Experiment 5.**

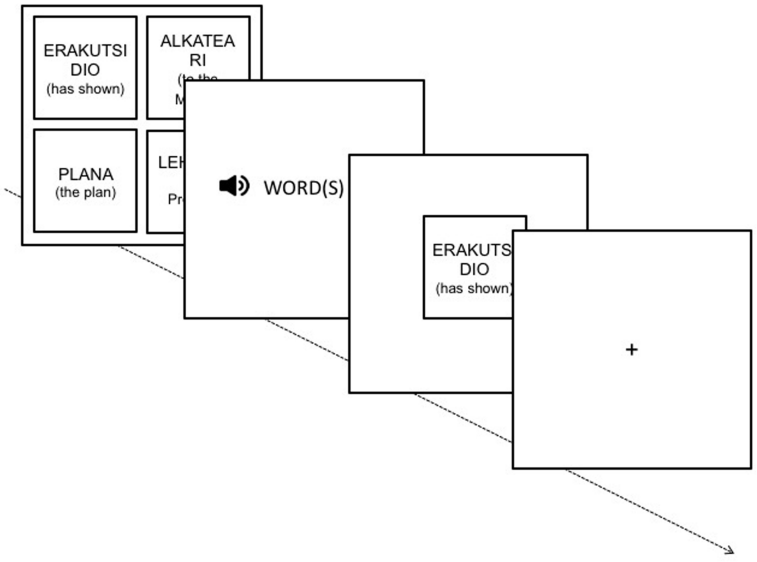


FIGURE 5.1: **Experimental procedure in the cue-based recall task in Experiment 5**

2.2.3 Scoring

A total of 130 (12.8%) responses were coded as miscellaneous and removed from further analyses in the same conditions as in Experiments 1-4 (see Section

2.3 2. session: Complex span tasks

During this second session, participants completed two automated complex span tasks, a Reading span task and an Operation span task, both created by Unsworth et al.(2005). They were translated to Basque by the author of this dissertation and are available through the Attention and working memory Lab of the Georgia Institute of Technology. This session lasted approximately between 45 and 60 minutes.

For an extremely thorough review on WM span tasks, from their genesis and current status in the field of psychology and others, to issues of reliability, validity and other technical aspects such as administration or scoring, see Cowan (2001). In contrast to simple span tasks such as digit or speaking span tasks, which do not require concurrent processing, in complex span tasks, the main recall task is interspersed with a relatively demanding secondary task. The nature of the task varies among complex span tasks and it can include digits, sentences, spatial locations, etc. The secondary tasks presented in this experiment consisted in reading and comprehending sentences in the Reading span task and in verifying arithmetic operations in the Operation span task. The goal of interleaving a secondary processing task is to be able to measure not only the ability of short-term storage, assessed by simple span tasks, but to tap in both the processing and storage parts of memory.

In this study, I used a continuous variable design and did not categorized scores to represent the upper and lower ends of the underlying distribution of WMC scores. Among other problems, extreme-group designs inflate effect sizes, add uncertainty about the relationship between the variables of interest across the full range of scores, and, with small sample sizes, statistical power to detect potentially interesting effects may be low, and confidence intervals large, allowing little confidence in replication of results. Also, subjects can be easily misclassified or treated as if they were equal when they are not (see Conway et al., 2005).

In this experiment, I included an Operation span task, and not only a Reading span task, since the verbal component of the processing task in the former is more limited than in the latter. The role of verbal WM on language use is still far from uncontroversial. Some researchers question the assumption that reading span tasks measure a system-external working memory capacity (Farmer, Misyak, and Christiansen, 2012; Van Dyke, Clinton, and Kukona, 2014; Wells et al., 2009; MacDonald and Christiansen 2002, Acheson et al, 2011, and Jones et al, 2015). In other words, reading span scores, which are overwhelmingly used in

much of the literature reviewed above (see also Daneman and Carpenter, 1980; Daneman and Merikle, 1986), might indirectly point to differences in linguistic experience. That is, the relationship reported above and scores in verbal WM span tasks could be due to shared variability in linguistic experience, and not to differences in intrinsic memory ability. In this experiment, if any effect is due to WM limitations and not to the effect of long-term linguistic representations on production preferences, one should be able to find it regardless of the span task used. If not, one should only find it when memory capacity is measured by the Reading span task.

2.3.1 Procedure

The Reading and the Operation span tasks were conducted in an individual sound-proof booth. They started with a practice session, which was broken down in three parts. In the first part, the participant was presented with a letter span. A letter appeared on the screen, and remained there for 800 ms. After that, it disappeared and another letter took its place. Participants saw two sets of 2 letters and 2 sets of 3 letters. After each set was presented, participants had to recall the letters in the order they were presented. They selected the letters from a 3 x 4 matrix where 12 letters where displayed (namely, F, H, J, K, L, N, P, Q, R, S, T, and Y), by clicking the box next to the appropriate letter in the required order. If they forgot any letter, they could fill in its position by pressing the blank button. When they finished, they clicked on the exit button and feedback was provided on the number of errors committed in that trial. In the original reading span task created by Danemann and Carpenter (1980), participants had to remember the last word of each of the 2-6 set of sentences that they had previously read out loud. Instead of words, in the two complex task I used participants had to recall letters (only consonants to hinder the use of any conscious strategies). This is so in order to reduce any influence of reading comprehension or verbal ability on the variance captured by the complex span task (see Conway et al., 2005 and references therein).

In the second part of the practice session, in the Reading span task, participants read sentences and checked their plausibility. The sentences presented in the experiment are reported in Appendix 1. Participants were presented with the sentences one by one. They saw one sentence and read it aloud. After that, they clicked to turn to the next screen, where they were required to decide whether the sentence made sense or not. In the Operation span task, simple arithmetic problems were used instead of sentences. Participants first saw a mathematical operation and were instructed to solve it as fast and accurately as possible. When they did, they clicked to advance to the next screen, where a digit was presented. Then, they had to decide whether the number that appeared on the screen was the correct response to the previously-seen operation. After each sentence or operation, participants were given accuracy feedback. The goal of this second part of the practice session in both the Reading and the Operation span task was to calculate how long it took each participant to read and decide about a sentence or to solve an arithmetic operation in order to account for initial individual

differences on the distracting task. The program calculated each individuals's mean for it. This time plus 2.5 standard deviation was used later during the experiment as time-out for solving the secondary task in each trial. In this part of the practice session, each participant saw a total of 15 math operations in the Operation span task and 15 sentences in the Reading span task-

In the third part of the practice session, participants saw the letters and the sentences together in the Reading span task and the letters and the arithmetic operations together in the Operation span task, as they did in the experimental trials (see Figure 5.2 for the Operation span task). Participants first saw the sentence or the arithmetic problem. Next, they clicked the mouse button to advance to the next screen when ready, where they had to decide whether the sentence was plausible or the arithmetic problem was correctly solved. If a participant took more time than her average time during the second part of the practice session plus 2.5 standard deviations, that trial counted as an error. The aim was to prevent any rehearsal strategy. In the practice session, participants completed 3 trials of set size 2. When participants finished the practice session, they proceeded to the experimental trials. These consisted of 3 sets each, with set sizes ranging from 3 to 7, always presented in a random order, so that participants could not anticipate the size of upcoming sets. In the real experiment, each participant saw 75 letters and 75 either arithmetic operations or sentences.

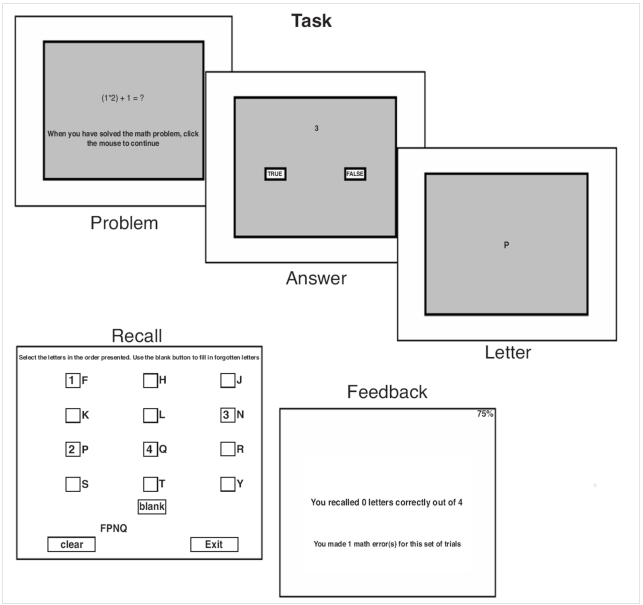


FIGURE 5.2: **Experimental procedure in the Operation span task** (image taken from Unsworth et al., 2005)

2.3.2 Scoring

To ensure that participants were attempting to solve both the processing (reading the sentences in the Reading span task and solving the operations in the Operation span task) and the storage parts of the task (recalling the letters), an 85% threshold for accuracy in the secondary task was used for all participants (see Conway et al., 2005). At the end of each trial, aside from feedback on the number of letters remembered, a percentage was presented in the upper right corner of the screen, to make participants aware of their performance in the secondary part of the task. They were encouraged to maintain that level at 85% or higher. As stated above in Section 2.1, the entire data set of 3 participants had to be removed from further analyses because they fell below this threshold in the processing component of the Operation span task. All participants scored above 85% in the Reading span task.

When scoring participants' responses, partial-credit unit scoring was used (following Conway et al., 2005). Unlike all-or-nothing scoring, in partial-credit scoring credit is given to partly correct items, and not to only those sets that have been completely correctly recalled. For example, in a set of letters of size 5, in an all-or-nothing scoring no points would be assigned when either 1, 2, 3 or 4 letters have been remembered. Also, I favoured unit-weight over load-weight scoring. In unit-weighting, all items within a set count the same independently of set size. In load-weighting, in contrast, items with a higher memory load (within a greater set), contribute more to the overall score. In sum, partial credit unit thus represents the mean proportion of elements within an item that were recalled correctly. As explained in Conway et al. (2005), load-weighted scoring is rarely used in psychometrics (although it is very much used when measuring span tasks), because all items are supposed to measure the same underlying ability. In this experiment, partial credit unit (PCU) scores ranged between 26 and 68 with an average of 47.9 (SD = 10.12) in the Reading span task and between 43 and 72 with a mean of 60 (SD = 6.7) in the Operation span task.

The correlation between the scores obtained by the participants in the Reading and in the Operation span task was marginally significant ($r(35) = .29$ [95% CI = -.03-.56], $p < .07$). Additionally, the number of words remembered in the production task was correlated with the results from the Reading Span ($r(35) = .34$ [95% CI = .03-61], $p < .04$), but not with the Operation Span task ($r(35) = .15$ [95% CI = -.17-.45], $p = .352$). When data from all subjects independently of their accuracy threshold were considered, all these correlations reached statistical significance ($r(45) = 35$ [95% CI = .07-.58], $p = .02$); $r(45) = .50$ [95% CI = .25-.69], $p < .0005$; and $r(45) = .34$ [95% CI = .05-.57], $P = .02$). The lack of shared variance between some of these tasks within the smaller sample is due probably to lack of statistical power.

2.4 Analysis

Shifted O-IO word orders, which yield shorter dependency lengths than their canonical counterparts, were regressed onto the main effects of constituent Length (All-Short vs. Long-O), Memory load (high vs. low), NP Position on the screen (consistent with shift vs. inconsistent with shift) and WMC, and their interactions as fixed effects. If the model did not converge, maximal random effects structures were reduced in a step-wise manner by removing the terms on the random effects structure to which the least amount of variance was attributed. Data was fitted to two different models. In the first one, WMC was measured via the Reading span task. In the second one, WMC was captured via the Operation span task. Both models included by-subject and by-item random slopes. Raw results are illustrated in Figure 5.3 (see also Table 5.2). Statistical analyses are reported in Tables 5.3 and 5.4. For the sake of simplicity, in the body of the text I report the results obtained in the first model (with the PCU scores obtained in the Reading span task as a fixed effect), since they do not differ in direction nor magnitude from the results from the second model (with the PCU scores obtained in the Operation span task as a fixed effect). A non-restricted analysis was also conducted, where I also included data from the subjects that did not exceed the accuracy threshold of 70% (n = 47 instead of 37). As in all previous experiments, the model contained the maximal random effects structure justified by the data based on backwards model comparison between all the models that converged and did not contain correlations between the random effects equal to 1/-1 (Baayen et al., 2008). I report p-values estimated by the R-package lmerTest (Kuznetsova et al., 2017). All analyses were conducted using R (R Core Team, 2014) and the lme4 package (D. M. Bates et al., 2014).

2.5 Results

2.5.1 Cue-based recall production task

In Experiment 5, participants tended to use canonical IO-O order overall more frequently than shifted orders ($\hat{\beta}$ = -1.043, z= -3.652, p < .0003): 67% of all the produced sentences were canonical. This is a 13% more than in Experiment 1 (see Section 2.7 in Chapter 2), where the effect of dependency length on word order preferences in Basque had been examined, regardless of the effect of memory load. With respect to the effect of the direct manipulation of Memory load on length-triggered word order preferences, an interaction emerged between Length, NP Position on the screen and Memory Load ($\hat{\beta}$ = .182, z = 2.009, p = .045). This interaction remained significant also when more simple models were used, without the inclusion of individual measures of WMC ($\hat{\beta}$ = .177, z = 1.966, p = .049). I next assessed the simple effect of the interaction of Memory Load and Length at each level of NP Position on the screen. Whereas I found a positive significant two-way interaction between Memory Load and Length when the position of the constituents on the screen was consistent with the

direction of shift (O-IO) ($\hat{\beta}$ = .400, z = 2.289, p= .022), there as no evidence for a corresponding interaction when the position of the constituents was inconsistent with the direction of shift (IO-O) ($\hat{\beta}$ = -.092, z = -.645, p= .519). As reported in Table 5.2 (see also Figure 5.3), if the order of the constituents on the screen was canonical (IO-O), when the O was long compared to when short, participants produced 7% more shifted orders if Memory load was low vs if high (Low: 20% vs. High: 13%). In contrast, if the order in which the constituents was consistent with the expected shifted order, that is O-IO, participants presented the reversed pattern. That is, when the O was long compared to when short, they produced shifted orders 8% more often when Memory load was high vs. low (Low: -13% vs. High: 5%).

		Constituent Length	
NP Position	**Memory load**	*All-Short*	*Long-O*
IO-O	Low	17 (12%)	48 (32%)
	High	20 (16%)	31 (29%)
Miscellaneous		15	42
O-IO	Low	60 (40%)	39 (27%)
	High	47 (36%)	42 (41%)
Miscellaneous		30	43

TABLE 5.2: **Raw data and percentages (in brackets) of shifted responses as a function of dependency length relative to memory load and position on the screen in Experiment 5.**

The main effects of Length, Memory load and NP Position on the screen that arose in my models were probably derived from the interaction reported above. When conditions were collapsed, participants placed the O before the IO only 6% more often when it was long (32%) than when short (26%) ($\hat{\beta}$ = .352, z = .3.813, p = .001), an 12% and a 16% less than in Experiments 1 and 4 respectively (see Section 2.7 in Chapter 2 and 3.5 in Chapter 4), where participants produced shifted *long-before-short* O-IO orders 18% and 22% more often when the O was long than when not. This difference across experiments is likely related to the effect of NP Position on the screen in this study. When all constituents were short, participants used 18% more shifted O-IO orders when this was the linearization previously seen on the screen (NP Position IO-O: 14% vs. NP Position O-IO: 38%). In contrast, when the O was long, there was virtually no difference between seeing the relevant constituents in an order consistent or inconsistent with shift (NP Position IO-O: 31% vs. NP Position O-IO: 33%) (see also Figure 5.3). As a consequence, the effect of the relative position of the constituents on the screen thus obscured any main effect of Length. Last, although the main effect of Memory load reached statistical significance ($\hat{\beta}$ = .281, z = 2.401, p = .02), participants only shifted constituent order overall when under a high concurrent memory load a 2% more often than when under a low memory load (Low: 28% vs. High: 30%).

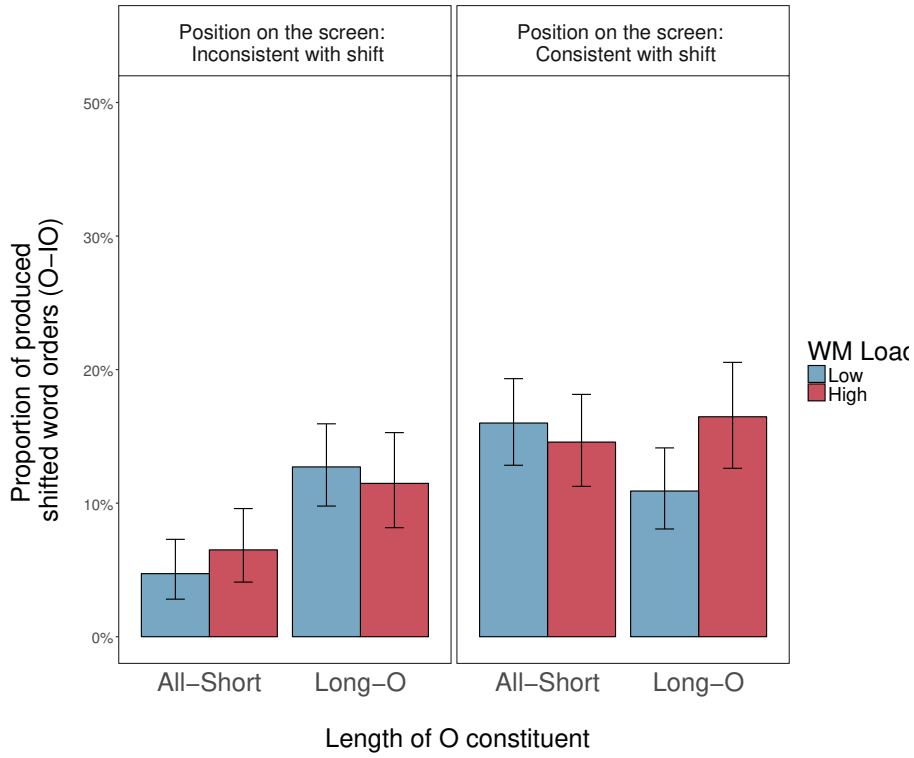


FIGURE 5.3: **Proportion of shifted responses produced in Experiment 5**. The error bars show 95% confidence intervals

In a non-restricted analysis, where subjects who fell below the 70% accuracy threshold (n = 47) were included, the main effects of Length and NP Position on the screen replicated ($\hat{\beta}$ = .423, z = 4.973, p < .0001 and $\hat{\beta}$ = .589, z = 5.427, p < .0001, when the scores in the Reading span task were included as a predictor; $\hat{\beta}$ = .415, z = 4.702, p <.001 and $\hat{\beta}$ = .613, z = 5.547, p = <.001, when the scores in the Operation span task were included as a predictor). However, the interaction between Memory load, NP Position and Length reported above did not reach significance in either model ($\hat{\beta}$ = .112, z = 1.361, p = .173 and $\hat{\beta}$ = .137, z = 1.598, p = .110, with Reading and Operation span scores as predictors, respectively)[6].

2.5.2 Complex span tasks

Regarding the effect of individual differences in WMC on length-related word order preferences, a significant interaction emerged between constituent Length and PCU scores in the Reading span task. When the O was long compared to when short, the higher the participant's scores in the complex span task, the greater the odds of resorting to O-IO shifted

[6]This results might be a statistical fluke or may be due to the fact that the discarded data comes from participants with lower memory spans. In fact, the number of errors committed (that is, the accuracy threshold) correlated with the PCU scores obtained in the complex span tasks (RSpan task: $\hat{\beta}$ = -1.737, df = 45, cor = -.25, p = 08 [95% CI: -.50-.03]; Operation Span task: $\hat{\beta}$ = -2.738, df = 45, cor = -.38, p = 009 [95% CI = -.50 -.04]).)

orders ($\hat{\beta}$ = .020, z = 2.136, p = .03) (see Figure 5.4). Recall that shifted O-IO orders yield shorter dependency lengths than their canonical IO-O counterparts. No significant interaction was found when WMC was measured by means of the Operation span task ($\hat{\beta}$ = .014, z =.905 , p =.366) (see Figure 5.5). In the non-restricted model mentioned in the previous subsection, where subjects who fell below the 70% accuracy threshold (n = 47) were included, the interaction between Reading Span and Length remained significant ($\hat{\beta}$ = .017, z = 2.114, p = .035).

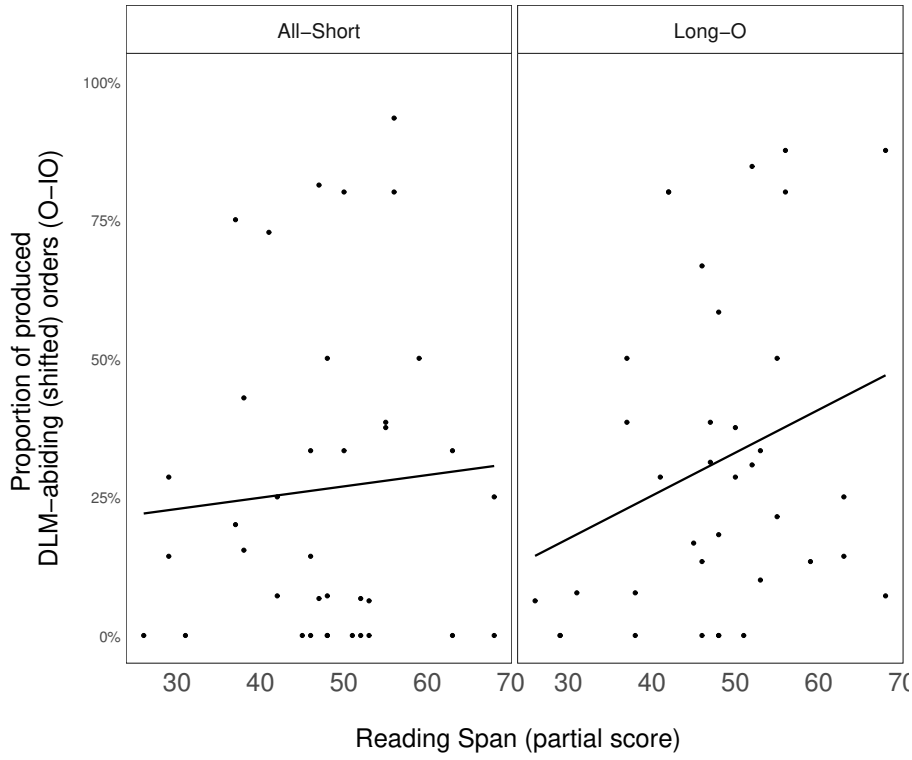


FIGURE 5.4: **Proportion of shifted responses by IDs in WMC as measured by a Reading Span task in Experiment 5**

In sum, in Experiment 5 speakers of Basque show an overall tendency to use canonical orders in ditransitive sentences, that was absent in a previous experiment in this language (see Chapter 2 and Ros et al., 2015). Also, the effect of dependency length was reduced in this experiment compared to previous ones (see Experiments 1 and 4 in Chapters 2 and 4), as the result of a stronger reliance on the position in which the constituents appeared on the computer screen. Crucially, the tendency to produce DLM-abiding *long-before-short* orders increases along WMC, as measured by a Reading span task, but not by an Operation span task. Also, the number of orders with shorter dependency length produced was modulated when under a concurrent memory load, but only when the O was presented to the left of the IO, in a O-IO reading order consistent with the expected direction of shift when reducing dependency length (namely, *long-before-short*). This effect was due to the fact that participants produced fewer O-IO

orders when under a low Memory load when the O was long than when short, whereas their number remained similar when under a high Memory load.

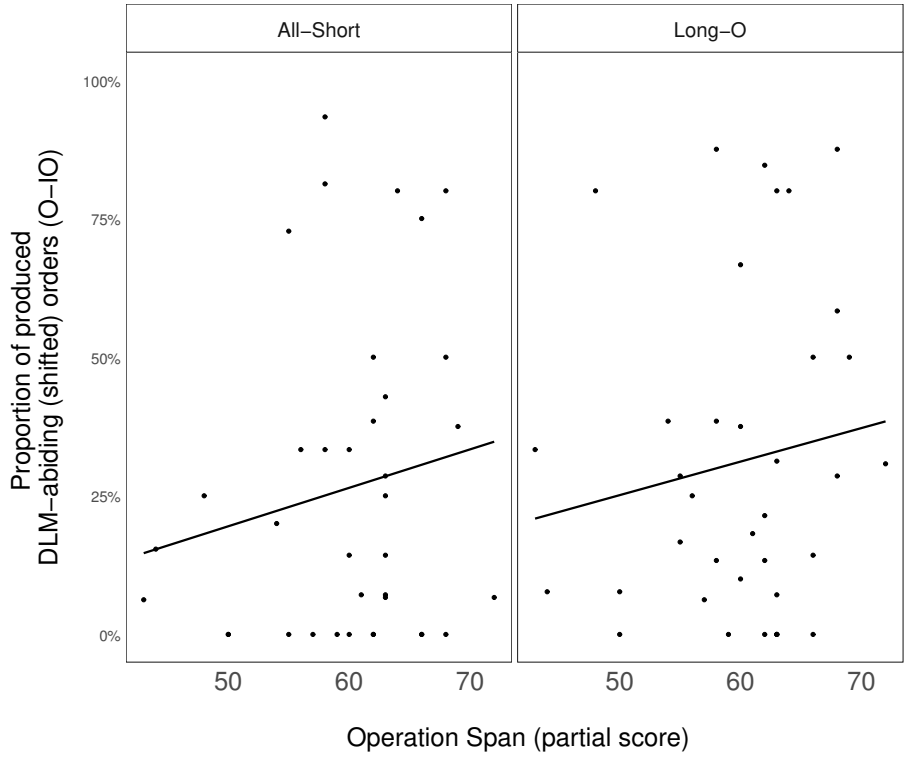


FIGURE 5.5: **Proportion of shifted responses by IDs in WMC as measured by a Reading Span task in Experiment 5**

	Estimate	**SE**	**z-value**	**p-value**
Total of sentences (N = 920)				
Intercept	-1.043	.286	-3.652	<.0003 ***
Memory Load	.218	.091	2.402	.02 *
Length	.352	.924	3.813	.001 ***
NP Position	.561	.109	5.166	<.001 ***
RSpan	.028	.028	1.007	.314
Memory Load * Length	.071	.113	.090	.431
Memory Load * NP Position	.076	.090	.842	.400
Length * NP Position	-.376	.278	-1.354	.176
Memory Load * RSpan	.002	.010	.179	.858
Length * RSpan	.020	.010	2.136	.033 *
NP Position * RSpan	.012	.010	1.240	.215
Memory Load * Length * NP Position	.182	.091	2.009	.045 *
Memory Load * Length * RSpan	.001	.009	.066	.948
Memory Load * NP Position * RSpan	.005	.009	.499	.618
Length * NP Position * RSpan	.029	.028	1.038	.299
Memory Load * Length * NP Position * RSpan	-.005	.010	-.516	.606

TABLE 5.3: **Logit mixed model analysis in Experiment 5, with the (partial) scores from the Reading Span span task as fixed effect.**

	Estimate	**SE**	**z-value**	**p-value**
Total of sentences (N = 920)				
Intercept	-1.060	.305	-3.481	.001 ***
Memory Load	.210	.0945	2.224	.026 *
Length	.363	.097	3.756	.001 ***
NP Position	.608	.113	5.383	.001 ***
OSpan	.030	.051	.591	.555
Memory Load * Length	.066	.094	.695	.487
Memory Load * NP Position	. 112	.094	1.188	.235
Length * NP Position	-.378	.297	-1.275	.202
Memory Load * OSpan	.018	.016	1.135	.256
Length * OSpan	.014	.016	.905	.366
NP Position * OSpan	-.007	.016	-.418	.676
Memory Load * Length * NP Position	.207	.095	2.188	.029 *
Memory Load * Length * OSpan	.012	.016	.818	.413
Memory Load * NP Position * OSpan	-.007	.016	-.432	.666
Length * NP Position * OSpan	.028	.051	.543	.587
Memory Load * Length * NP Position * OSpan	-.002	.016	-.104	.917

TABLE 5.4: **Logit mixed model analysis in Experiment 5, with the (partial) scores from the Operation Span span task as fixed effect.**

3 Discussion

In this study, I explored the effects of memory demands on the bias to reduce dependency length in production data. In order to do so, a cued-recall dual-task was conducted, and participants' WMC was assessed via a Reading and an Operation complex span task. I found that the preference for producing shifted word orders that yield shorter dependency lengths than their canonical counterparts increases along with WMC, as measured by a Reading span task. Parallel results have been reported in comprehension, as reported in Section 1.2, where readers with higher WMC, compared to those with lower WMC, seemed to be more sensitive to any difficulty derived from the presence of non-local dependencies in the input (Nicenboim et al., 2016; Swets et al., 2007). However, I failed to find any relation of word order preferences and WMC when it was measured via an Operation span task. Additionally, compared to previous experiments where DLM in Basque was investigated (see Chapters 2 and 4), participants in this study produced an overall higher rate of canonical word orders. Last, when the relevant constituents were presented in a left-to-right reading order consistent with the expected direction of shift in Basque, participants tended to produce *long-before-short* orders less often when under a low Memory load than under a high Memory load. In sum, the results of this experiment contrast with the idea that DLM is caused by memory limitations (Futrell, 2017; Liu, 2008; Scontras et al., 2015) and align better with the claim of information-theoretic approaches to language form that DLM is derived from considerations of communicative efficiency (Fedzechkina et al., 2018; Ferrer i Cancho, 2004; Futrell et al., 2015b; Gildea and Jaeger, 2015; Hawkins, 2014). I first discuss the implications of these results for a framework that understands communication as derived from rational considerations of efficiency. After that, I tackle some plausible alternative explanations for some of the findings obtained in this study. Specifically, I will address the possible roles of planning scope and of verbal ability with regard to the preference for short dependencies.

3.1 Efficient communication and the preference for DLM

The positive interaction between verbal WMC and the preference for DLM cannot be straightforwardly accommodated by the claim that the bias for short dependencies stems, as in comprehension, from limitations in human memory, either due to a hindered ability to complete a retrieval or to maintain elements in the focus of attention, to a reduced capacity in allocation of attention or to upper limits on storage capacity. The greater the memory ability of the participants, the more often they resort to DLM. This novel finding reinforces the hypothesis that the bias for DLM is the result of the need to reduce the cognitive cost associated with producing shifted word orders while assuring the probability of correctly retrieving the intended message. In Chapters 1 and 3, I argued that shifted orders are more costly than their canonical counterparts due to the fact that alternative candidates are activated simultaneously when producing them. At the same time, when non-local

dependencies are involved, the odds of communicative success are increased by reducing dependency length whenever possible, since short dependencies have been shown to be more easily processed. If producing shifted word orders is costly, it straightforwardly follows that low-WMC speakers will produce them to a lesser extent than their high-WMC counterparts, at the expense of successful transmission of the intended message. Speakers with higher memory spans should be able to maintain alternative structural candidates in mind more easily, and choose the most appropriate one for their communicative intentions, in the same vein as King and Just (1991).

Regarding the results in the cued-recall dual-task, I found evidence for a weaker effect of DLM when Memory load is low versus high when participants were previously presented with O and IO in an order consistent with the expected direction of shift (O-IO). Nevertheless, any possible interpretation of the interaction between the variables of Length and Memory load is hindered because of the effect of NP Position on the screen. NP Position on the screen affected shifted preference to a greater extent when both constituents involved were short than when O was long. When constituents were all short, participants used shifted orders a 24% more when O and IO were previously seen in an O-IO order. When the O was long, participants used shifted orders only a 3.5% more when they were previously seen in an O-IO order. The effect of NP Position on the screen thus obscures the effect of memory demands in the variable of interest, namely the proportion of shifted orders when a long versus a short constituent is involved. Also, I acknowledge that it is possible that any lack of reliable results in the experiment is due to the fact that, although most of the to-be-recalled words were more similar to the O than to IO insofar as the were inanimate, some of them were animate as the IO (see Section 2.2.1), which could have added a certain level of noise to the data.

Participants also shifted orders overall less than in previous experiments where I tested the preference for DLM in Basque, where Memory load was not manipulated (see Chapters 2 and 4 and Ros et al., 2015). This is also in line with the idea that shifted orders are somehow cognitively more costly to produce than their canonical counterparts. My experimental task imposed greater cognitive demands than previous ones: Participants had a time limit to read the unordered sentence fragments and prepare the sentence and they had to remember either one or three words before they could start uttering the intended sentence. Verbatim sentence recall, as in most of the sentences produced by participants in this experiments, has been claimed to be an attentionally demanding verbal task (Schweppe, Barth, Ketzer-Noltge, and Rummer, 2015), and its cognitive cost could be exacerbated when also asked to remember some extra words. As MacDonald (2013) argues, when faced with an especially stringent production task, speakers could tend to reuse more frequent sentence plants saved in long-term memory.

In sum, initial evidence from this study is congruent with the hypothesis that DLM is the outcome of efficient processing. Replication of my results is necessary, and it is of paramount importance to choose an experimental design where any possible priming effects of word

orders previously seen on the screen could canceled out. Also, it would be desirable to reduce the difficulty of the task, so that the range of memory spans included in the analysis could be broader, and to more adequately control for the type of stimuli used in the Memory load condition.

3.2 DLM and the scope of language planning

Swets et al. (2007) found that WMC is inversely related to the preference for local attachment site in ambiguous relative clauses. To account for their seemingly counterintutive results within the framework of the assumptions of memory-based theories of language processing, Swets and colleagues argued that memory capacity could impact the size of segments of text that participants can maintain in memory during silent reading. This, in turn, indirectly modulates speakers' preference with regard to local vs .non-local attachment preferences. Similarly, in production, it has been shown that the size of planning scope is flexible (for a review, see Konopka and Brown-Schmidt, 2014). Swets et al. (2014) found that the size of planning scope is broader for speaker with higher than with lower memory spans. If this is so, WCM could very likely affect the bias for short dependencies only spuriously, due to its relationship with planning scope strategies.

Evidence from corpus and experimental studies suggests that relative and not absolute length is the key factor behind the magnitude of the preference for producing shifted orders that minimize dependency length (Hawkins, 1994; Stallings and MacDonald, 2011; Wasow, 2002). Stallings and MacDonald (2011) constructed 30 items containing an object NP of 10 words and a PP of 2, 5, or 7 words. Participants shifted reliably more in the short (2-word) PP condition (38.7%) than in the long (7-word) condition, even when the length of the NP remained identical in both cases (10-word). 5-word PP condition shifting did not significantly differ from either of them. In a second experiment, two conditions were matched for relative length (5 words), but varied with respect to absolute length (Condition 1: NP = 10 words, PP = 5 words; Condition 2: NP = 7 words, PP = 2 words). They found no effect on ordering relative to absolute length. Condition 1, with the same 5-word difference as condition 2 but a 3-word longer NP difference, did not yield more shifting than condition 2. If the length of both constituents involved in shifting is the key factor in determining word order preferences, the size of planning scope required for reducing dependency length should be quite broad. It is thus possible that high-WMC individuals are more apt at reducing dependency length because, when producing language, they can rely on broader planning scope strategies more easily than low-WMC individuals. This hypothesis opens up the possibility that producing shifted orders that yield shorter dependency lengths is costly not only because alternative orders need to compete with each other, but because the scope of language planning required when shifting constituents places larger demands on memory capacity.

3.3 Working memory or linguistic ability?

As explained in Section 2.3, an Operation span task was included in the set of experiments included in this study due to the current debate over whether reading span tasks reliably measure WMC or are better regarded as a proxy for verbal capacity, which is informed by differences in linguistic experience among individuals across their lifespan. Farmer et al. (2017) found a positive correlation between proxy measures of linguistic experience and a reading verbal span task, but not with a non-verbal Backward Digit Span simple (or storage-only) test, while also reporting a positive correlation between the strength of the garden-path effect and the verbal WM task. Others, in contrast, have found that processing of non-local dependencies interact both with verbal and arithmetic integration processes (Fedorenko et al., 2007) or that are correlated with WM as measured by an Operation span task similar to the one used in my study (Nicenboim et al., 2016; Nicenboim et al., 2015).

As reported above (see Section 2.5.2), I failed to find any evidence of any relation between WMC as measured by an Operation span task and length-triggered word order preferences. Turner and Engle (1989) created the Operation span task specifically to be less sensitive to any possible influence of language ability. That is, the main difference between the Operation and the Reading span tasks used in this experiment is the amount of language processing required by each. Therefore, the results in Experiment 5 do not preclude the possibility that WMC could be only spuriously related to DLM[7]. It is thus possible that the bias towards DLM could be connected to verbal or linguistic skills and not to intrinsic WMC. This hypothesis is in line with the idea that DLM could be modulated by individual and strategic differences in the size of planning scope chunks delineated in the section above. With regard to the relation between planning scope and DLM, Christiansen and Chater (Christiansen and Chater, 2016) have proposed that linguistic experience might facilitate chunking and subsequent processing and integration strategies of chunked input. In any case, he role of linguistic experience in modulating DLM preferences in production deserves further scrutiny.

4 Concluding remarks

The cognitive bias for short dependencies has been attested in a wide number of languages in both corpus and experimental studies, suggesting that this principle is a cornerstone property of human language, that impacts both grammars and usage preferences. Still, the underlying causes for this fundamental principle are heatedly debated in research on language comprehension and have been so far neglected in the field of language production. Both the nature of the human memory system as a limited-resource and of language as an efficient

[7]The lack or absence of an effect of WMC might also be due to lack of statistical power. No previous study on the role of WM on DLM has been previously conducted and no estimate of effect size was available to calculate the power needed to reduce Type II or S/M errors

communicative system have been proposed as possible sources of the tendency to reduce dependency length. In the production study reported in this chapter, I intended to provide a first glimpse of how memory demands and memory capacity interact with the preference to minimize dependency length.

Throughout this dissertation, I have argued that DLM is the outcome of one main architectural specifications of the linguistic system, its efficiency. Language is a system highly tuned to our communication needs, shaped by the goal to robustly convey any intended meaning, with the least cognitive effort this could entail. DLM responds to this requirement of efficiency, insofar as, in spite of entailing a certain cognitive cost, it is in line with listeners' preference for short dependencies. My results converge with this idea. First, I do not find evidence that more stringent memory demands enhance the preference for reducing dependency length. Second, the bias towards short dependencies increases alongside WMC. These findings are in contrast with the idea that DLM in production is the outcome of memory demands imposed by long dependencies. As in comprehension, the interplay between linguistic and processing preferences and memory demands could be less clear-cut than originally assumed in theories that conceive of memory as a resource bounded by its upper limits. Further assumptions would be needed to accommodate the positive correlation between WMC and the preference for short dependencies, which points to a less straightforward linking hypothesis between memory and word order preferences.

My findings contribute a first data point in the approximation to the cognitive causes of word order preferences in production as a function of dependency length. Doubtless, further research is needed to replicate my results in the face of some possible shortcomings of my original design. But, in any case, the results reported in this chapter pose a set of new research questions regarding the nature of the relation between memory and DLM, whether intrinsic properties of the human memory system are only spuriously or indirectly related to the bias towards short dependencies, and with respect to the role of the scope of language planning and of verbal ability in the emergence of the bias towards short dependencies.

5 Summary of findings

1. DLM is modulated by WMC as measured by a Reading complex span task. The preference for short dependencies increases alongside WMC. This is in contrast with any theoretical model that proposes that DLM is the result of the higher demands imposed by long dependencies to memory and provides indirect support to the idea that the production of shifted orders is costly, a key assumption in the hypothesis that DLM stems from a trade-off between increasing redundancy in the signal and decreasing production effort.
2. I find no evidence of any modulation of WMC and DLM, when memory capacity is assessed using an Operation span task. This null result calls into question any

explanation that sidelines verbal ability as a possible factor in modulating the tendency to reduce dependency length, and, together with the high reliance on written presentation of the relative order of the constituents involved in shifting, underscores the need for further research on the connection between production and comprehension processes.

3. In a cued-recall dual-task production experiment, participants resorted to DLM more often when under a low than when under a high Memory load, but only when they had previously seen the constituents relevant for shifting in a written order consistent with the expected direction of shift in Basque. This result suggests that there might be a certain relation between memory demands and dependency length, but I refrain from any possible interpretation due to the effect of the order of presentation of the constituents on the screen. Further research where this factor can be eliminated is necessary to fully address this topic.

Chapter 6

General conclusions and questions for further research

The main contribution of this dissertation, where I investigated production preferences for short dependencies across languages and speakers, is three-fold:

1. I showed that the bias for reducing dependency length is present in production preferences in languages as diverse as Basque, Polish and Spanish (see Chapters 2 and 3). This constitutes strong evidence that DLM is a general principle active in human language processing.

2. I provided unprecedented evidence that variation in the bias for reducing dependency length is present across and within languages (see Chapter 3) and across speakers (see Chapters 4 and 5). I sought to reconcile this finding with information-theoretic considerations, and suggested that near-optimal communicative efficiency can be obtained across languages through several mechanisms:

   (a) When variation is found across languages, I highlighted the role played by language-specific properties like pluripersonal agreement and by cognitive-general mechanisms like building input-derived expectations, and addressed how these two principles could be intertwined with the preference for reducing dependency length (see Chapters 2 and 3, respectively).

   (b) When variation is found across speakers, I claimed that it can be accounted for by differences in the degree of language competence and in working memory capacity (see Chapters 4 and 5, respectively).

3. I explored the role played by the architectural properties of the human memory system in modulating the bias towards DLM (Chapter 5). I failed to find any evidence to support the view that short dependencies are due to working memory limitations. My results align better with the idea that the bias towards short dependencies is the result of communicative efficiency-related considerations.

In Sections 1 to 3 in this chapter, I survey the major findings of this dissertation, outlined in the previous paragraphs, and discuss their implications. In Section 4, I address some remaining questions and comment on some possible extensions to this work.

1 Cross-linguistic evidence for DLM

One main claim put forth in this dissertation and in previous literature is that DLM is grounded in general principles of human cognition. It is thus clear that cross-linguistic research is extremely relevant to test this claim. In this respect, the experiments in Part 1 of this dissertation, where typologically diverse languages are investigated, further contribute to currently growing evidence supporting the claim that the general bias towards reducing dependency distance is a major determinant of production preferences. Also, cross-linguistic research as the one carried out in this work helps us point out unadverted gaps in the theoretical coverage of this specific phenomenon (see Section 4 below).

Evidence from Experiment 1 in Basque is vital to support the hypothesis that DLM is a general processing bias grounded in human cognition (see Chapter 2). The mirror-like *short-before-long* and *long-before-short* ordering of constituents of different length present in several languages has been related to the differential effect of accessibility due to language-specific properties in rigidly head-final languages (Chang, 2009; Yamashita and Chang, 2001). Furthermore, the *long-before-short* preference has been attributed to the tendency to avoid center-embeddings in languages where syntactic constituents cannot appear after the verb (Wasow, 2002, 2013). In Chapter 2, I find a *long-before-short* preference when ordering constituents of different length in Basque, a head-final language that allows postverbal constituents (De Rijk, 1969; Hualde and Ortiz de Urbina, 2003). This constitutes strong evidence in favour of DLM as the result of a principle towards efficient processing. Furthermore, in Experiment 2, I replicated the preference for a *short-before-long* linearization of constituents of different length in two head-initial languages other than English, namely in Spanish and Polish, in an experimentally controlled setting (see Chapter 3). The importance of direct replications in psychological science cannot be sufficiently underscored as the cornerstone of scientific advancement, especially when the field of psychology (among many others) has been confronting serious questions about its research practices and the reproducibility of certain key findings (Open Science Collaboration and others, 2015).

Summing up, evidence from Basque, Polish and Spanish support the claim that there is a production bias towards those word orders that yield shorter dependency lengths, when more than one possible alternative is available to convey the same meaning. My research shows that the opposite linearization preferences employed to minimize dependency length in head-final and head-initial languages, originally interpreted as the result of language-specific differences

with regard to accessibility effects, points to a deeper cross-linguistic generalization, namely a general cognitive tendency towards DLM.

2 Language-specific variation in the bias for DLM

The experiments reported in Part I of this dissertation show that languages vary with regard to how they achieve communicative efficiency and how strongly they rely on DLM to do so. In Chapter 2, I showed that there is variation in how languages facilitate processing depending on grammatical properties live verb agreement (Experiment 1). In Chapter 3, I uncovered differences in the extent to which typologycally diverse languages rely on DLM (Experiment 2).

In Chapter 2, I provided evidence that efficient language processing can be further enhanced by pluripersonal verb agreement. When constituents are shifted to reduce dependency length, verb-medial orders tend to be used more frequently. I argue that, when faced with less frequent linearizations, rich agreement acts as a reliable cue for comprehenders by signaling upcoming constituents, and thus decrease their cost of integration when encountered. Additionally, verb-medial orders can reduce the space of possible continuations in the speech signal and thus the overall entropy with respect to the number of upcoming linearizations available to the speaker and the comprehender (Dye, Milin, Futrell, and Ramscar, 2018).

In Chapter 3, I found consistent variation in the strength of the preference for DLM in Basque, Polish and Spanish. Specifically, I explored the role of head position and case marking in the strength of the preference for DLM. I found that languages with rich case marking systems do not conform with a weaker reliance on DLM. These results are not compatible with the view that case marking adds redundancy to the speech signal and increase the probability of successful communication (MacWhinney et al., 1984; Sasaki and MacWhinney, 2006; Tily, 2010). In Chapter 3, I argued that word order-based distributional properties can account for differences in the magnitude of the bias for DLM. More precisely, expectations about upcoming linguistic materials stemming from language-specific distributions alleviate the cost of processing long dependencies in comprehension, which in turn would modulate the need to resort to DLM in production, if speakers abide by considerations of efficiency.

Summing up, this dissertation clearly shows that languages react differently with respect to the tendency towards DLM, and that this variation is dependent on word order. I argue that intra- and cross-linguistic variation in the bias towards DLM is compatible with an information-theoretic perspective of language form. This frees us from having to postulate differences in communicative efficiency across languages. Nevertheless, as pointed out in Chapter 3, DLM in head-initial languages is correlated with information structure

(crosslinguistically, *given* elements are placed before *new* ones, and tend to be *short*, see Arnold et al., 2000; Wasow, 2002, among others). Also, word order freedom in the type of structures employed in my experimental design (transtives and ditranstives) is greater in Basque than in Polish and Spanish, which could have interacted with the preference for DLM across these languages (although we are lacking an objective measure of word order freedom across languages, see Futrell, Mahowald, and Gibson, 2015c; Gulordava and Merlo, 2016, *inter alia* for promising research in this direction). The work presented in this dissertations is thus a point of departure for further research on the casual pathways of the interaction between language-specific properties and the universal preference for short dependencies.

3 Individual variation in the bias for DLM

In Part 2 of this dissertation, I identified two factors that modulate the preference for DLM across speakers: nativeness (Experiments 3 and 4) and working memory capacity (Experiment 5). In Chapter 5, I explored length-triggered word order preferences in native versus non-native bilingual speakers of Spanish and Basque, a head-initial and a head-final language which present mirror-like *short-before-long* and *long-before-short* word order preferences. I tested the hypothesis that non-native highly proficient bilinguals transfer the superficial linearization favored in their native language to their non-native one (Dennison, 2008). Also, I sought to test the anecdotal observation that non-native speakers tended to use Basque in a more demanding way than native speakers, from a listener's perspective (Maia, 2015; Zubimendi and Esnal, 1993, *inter alia*). I showed that highly proficient and early non-native bilinguals of both Basque and Spanish reduce dependency length less often than their native peers, either monolinguals or bilinguals. This suggests that even early and highly proficient non-native bilinguals use language less efficiently than natives, since they use long dependencies more frequently, which are more costly for the listener (Gibson, 2000; Grodner and Gibson, 2005, among others). Also, I found no evidence for cross-linguistic transfer of the superficial *short-before-long* and *long-before-short* orderings present in the native language of the bilingual participants to the non-native one, contrary to previous claims (Dennison, 2008). In sum, the results from Experiments 3 and 4 reported in Chapter 5 indicate that non-native speakers unbalance the trade-off between production cost and successful transmission of information in favor of the former.

In Chapter 6, I explored the underlying causes for the preference for DLM in production. Several production studies have suggested that short dependencies are less taxing in WM than longer ones (Liu, 2008; Scontras et al., 2015; Temperley and Gidea, 2018, among others). I did not find any evidence supporting this hypothesis. Moreover, particular upper limits in WM have been directly linked to measures of average dependency lengths (see Liu, 2008). Dependency length measures are proxies for text or sentence complexity and stand for yet unkown cognitive computations involved in the processing of long dependencies. In this

sense, it is crucial to make as few assumptions as possible about them. Furthermore, any connection between memory processes and language processing needs to be critically and experimentally examined. First, because theoretical assumptions appropriate for comprehension have been rashly adopted in language production, without enough empirical evidence (see Chapters 1 and 5). Second, because some claims regarding the constraints imposed by human memory limitations are built upon what has been proved to be unwarranted assumptions regarding the complexity of certain cognitive computations (Jaeger, 2013; Tanenhaus, 2004, 2013). The results in Experiment 5 rather suggest that communicative efficiency is a more promising avenue for future inquiry (Fedzechkina et al., 2018; Ferrer i Cancho, 2005; Hawkins, 1994, among others), since they suggest that WM is negatively correlated with the production of shifted orders, a result fully compatible with the idea that DLM is costly for speakers. Additionally, the results from Experiment 5 underscore the presence of individual variation with respect to the preference for short dependencies in production and the need to take into account verbal ability as a possible factor in modulating it.

4 Directions for further research

One main claim in this dissertation and in much of the literature cited therein is that minimizing dependency length when speaking is communicatively efficient because this tendency agrees with a universal bias towards locality in comprehension (Bartek et al., 2011; Demberg and Keller, 2008; Gibson, 2000; Grodner and Gibson, 2005; Levy et al., 2013, among others). It is clear, however, that languages are not optimally efficient, that is, they do not reach the highest possible ratio of DLM allowed by their grammars (Futrell et al., 2015b; Gildea and Jaeger, 2015). This is to be expected in a complex system as language, where word order preferences have been shown to require a multifactorial approximation (Bresnan et al., 2007; MacDonald, 1994; MacDonald et al., 1994, among others). The effects of accessibility, the use of other strategies besides DLM to obtain short dependencies, as well as the role of pragmatic considerations, among other aspects, all have an impact on word order preferences in production. For instance, in head-final Basque, where length and information structure effects are not correlated, the bias towards DLM is lower than in head-initial Spanish and Polish, where they are (*given* elements tend to be *short*) (see Chapter 3). If, in certain languages, there is only a minimal production preference for more efficient word orders with short dependencies, it is unclear how DLM can come to affect their grammars in such a profound way as claimed in the literature. Further research is needed to obtain a more reliable estimate of the preference for DLM in language use. Work aiming at disentangling the effect of information structure and DLM in head-initial languages is needed. Also, the role of other grammatical properties like word order freedom in modulating DLM needs to be addressed. A clear picture of the nature and extent of usage preferences regarding DLM is an essential first link in the chain of hypotheses that lead up to positing DLM as a universal principle

shaping grammars and language change (see Section 1 in Chapter 1). The work described in this dissertation is a first step in that direction. In this respect, even if the bias towards DLM and, more generally, communicative efficiency influence adult language use, it is imperative to investigate whether and how they guide language acquisition.

It is possible that divergences in DLM thresholds across and within languages will still arise, even when factors such as distributional properties or information packaging mechanisms are factored out. This could indicate that some languages are more efficient than others, which may be of special interest for researchers investigating language change. An alternative possibility is that efficiency ratios as measured by average dependency lengths are not linearly correlated with processing difficulty, or that the nature of this relation varies depending on the make-up of each specific language. In this respect, *anti-locality* effects (facilitation when long dependencies are involved) have been reported in several head-final languages such as Hindi, German, or Persian (Konieczny and Döring, 2003; Nakatani and Gibson, 2010; Vasishth and Drenhaus, 2011, among others). There is no current data available in Basque, Polish nor Spanish. Further work is needed to fill this gap by investigating how tolerable long dependencies are in comprehension in these languages, and how strongly behavioral signatures of the cognitive cost of long dependencies correlate with each other in production and comprehension. The type of research advocated for in this and the previous paragraph can also be informative with respect to the effect of efficiency requirements in the case of non-native speakers of languages with opposing head positions. Once upper and lower bounds in average dependency lengths across languages and their effect on listeners' behaviors are better known, they can be used as reliable baselines with which we can compare non-native speakers' compliance with the preference for DLM. Also, further research on late and less competent bilinguals would allow to address the issue of whether there are qualitative changes across bilingual speakers' developmental trajectories.

Last, further work is necessary on the role of human memory architectural properties in the emergence of the bias for short dependencies in language production. The study reported in Chapter 6 is a first step studying the underlying causes of DLM. Further research is of paramount importance in addressing the role of both memory and communicative requirements in the preference for short dependencies. Also, I found systematic and considerable variation across individuals with respect to this preference. Some of this variation can be explained away by the effect of individual differences in verbal WMC. Nevertheless, most of it remains unexplained. This variation complicates any plausible attempt to capture how a processing constraint such as DLM could affect grammars and language change, since in most iterative experiments and simulations on the effect of processing constrains and learning on language form, there is only one direct and consistent source of input for the learners (Culbertson et al., 2012; Kirby et al., 2008; Kirby, Tamariz, Cornish, and Smith, 2015, *inter alia*). Further research needs to include variability as an aspect to account for when drawing a causal connection between processing preferences and grammatical properties across languages .

This dissertation has focused on investigating sentence word order production preferences as a function of dependency length across different languages, native and non-native bilingual speakers and speakers with variable working memory capacity. Though belonging to the area of psycholinguistics, I believe my work can be of interest to other fields such as Natural Language Processing (NLP) and linguistic typology and further encourage already on-going interdisciplinary cross-pollination. Gomez Rodríguez (2017), for instance, underscores the parallelism between some of the language principles claimed to be the result of general cognitive constraints and the solutions brought by NLP to the task of designing accurate and efficient parsing systems (see Gómez-Rodríguez, 2016 and Christiansen and Chater, 2016), such as the increase in parsing performance when short dependencies are involved (see Merlo, 2015 and references therein). Regarding linguistic typology, given the relatively easy access to large databases and the existence of appropriate statistical methods, benchmark findings in psycholinguistics can be employed as realiable tools for typologists trying to connect distributional patterns across languages in the world and general principles grounded in human cognition (see Bickel, 2015; Jaeger and Tily, 2010). Similarly, attested cross-linguistic distributional patterns can inform new psycholinguistic theories and testable hypotheses.

Appendix A

Appendices to Chapter 1

1 Linguistic Questionnaire used in Experiments 1-5

Personal information		
First name:		
Last name:		
Date of birth:		
Sex	Male	Female
E-mail:		
Telephone:		
Place of birth:		
Place of residence:		
Have you ever lived in another place?	Where?	When?

Age of acquisition	
How old were you when you started...	
speaking Basque?	
speaking Spanish?	

First language	
When your were little, what language did you use to use with...	
your mother?	
Basque	1
Mostly Basque, but sometimes Spanish	2
Both Basque and Spanish	3
Mostly Spanish, but sometimes Basque	4
Spanish	5
your father?	
Basque	1
Mostly Basque, but sometimes Spanish	2
Both Basque and Spanish	3
Moslty Spanish, but sometimes Basque	4
Spanish	5
your siblings?	
Basque	1
Mostly Basque, but sometimes Spanish	2
Both Basque and Spanish	3
Mostly Spanish, but sometimes Basque	4
Spanish	5
your grandparents?	
Basque	1
Mostly Basque, but sometimes Spanish	2
Both Basque and Spanish	3
Mostly Spanish, but sometimes Basque	4
Spanish	5

Language use	
Which language and how often did you use it...	
when you were little, before going to school?	
only Basque	1
mostly Basque, only sometimes Spanish	2
mostly Basque, but Spanish at least 25%	3
Basque and Spanish with the same frequency	4
Mostly Spanish, but using Basque at least 25%	5
Mostly Spanish, but sometimes Basque	6
only Spanish	7

	In school/ at university/ at work		***At home***		***Other places***	
in school?	only Basque	1	only Basque	1	only Basque	1
	mostly Basque, only sometimes Spanish	2	mostly Basque, only sometimes Spanish	2	mostly Basque, only sometimes Spanish	2
	mostly Basque, but Spanish at least 25%	3	mostly Basque, but Spanish at least 25%	3	mostly Basque, but Spanish at least 25%	3
	Basque and Spanish with the same frequency	4	Basque and Spanish with the same frequency	4	Basque and Spanish with the same frequency	4
	Mostly Spanish, but using Basuqe at least 25%	5	Mostly Spanish, but using Basuqe at least 25%	5	Mostly Spanish, but using Basuqe at least 25%	5
	Mostly Spanish, only sometimes Basque	6	Mostly Spanish, only sometimes Basque	6	Mostly Spanish, only sometimes Basque	6
	only Spanish	7	only Spanish	7	only Spanish	7
in high school?	only Basque	1	only Basque	1	only Basque	1
	mostly Basque, only sometimes Spanish	2	mostly Basque, only sometimes Spanish	2	mostly Basque, only sometimes Spanish	2
	mostly Basque, but Spanish at least 25%	3	mostly Basque, but Spanish at least 25%	3	mostly Basque, but Spanish at least 25%	3
	Basque and Spanish with the same frequency	4	Basque and Spanish with the same frequency	4	Basque and Spanish with the same frequency	4
	Mostly Spanish, but using Basque at least 25%	5	Mostly Spanish, but using Basuqe at least 25%	5	Mostly Spanish, but using Basque at least 25%	5
	Mostly Spanish, only sometimes Basque	6	Mostly Spanish, only sometimes Basque	6	Mostly Spanish, only sometimes Basque	6
	only Spanish	7	only Spanish	7	only Spanish	7
nowadays?	only Basque	1	only Basque	1	only Basque	1
	mostly Basque, only sometimes Spanish	2	mostly Basque, only sometimes Spanish	2	mostly Basque, only sometimes Spanish	2
	mostly Basque, but Spanish at least 25%	3	mostly Basque, but Spanish at least 25%	3	mostly Basque, but Spanish at least 25%	3
	Basque and Spanish with the same frequency	4	Basque and Spanish with the same frequency	4	Basque and Spanish with the same frequency	4
	Mostly Spanish, but using Basque at least 25%	5	Mostly Spanish, but using Basuqe at least 25%	5	Mostly Spanish, but using Basque at least 25%	5
	Mostly Spanish, only sometimes Basque	6	Mostly Spanish, only sometimes Basque	6	Mostly Spanish, only sometimes Basque	6
	only Spanish	7	only Spanish	7	only Spanish	7

Proficiency: self-assessment									
Rank the following skills:									
Basque									
	Speaking		Listening		Reading		Writing		Certificate
	very poorly	7	very poorly	7	very poorly	7	very poorly	7	
	poorly	6	poorly	6	poorly	6	poorly	6	
	justu	5	justu	5	justu	5	justu	5	
	enough	4	enough	4	enough	4	enough	4	
	well	3	well	3	well	3	well	3	
	very well	2	very well	2	very well	2	very well	2	
	perfect	1	perfect	1	perfect	1	perfect	1	
Spanish									
	Speaking		Listening		Reading		Writing		Certificate
	very poorly	7	very poorly	7	very poorly	7	very poorly	7	
	poorly	6	poorly	6	poorly	6	poorly	6	
	justu	5	justu	5	justu	5	justu	5	
	enough	4	enough	4	enough	4	enough	4	
	well	3	well	3	well	3	well	3	
	very well	2	very well	2	very well	2	very well	2	
	perfect	1	perfect	1	perfect	1	perfect	1	
English									
	Speaking		Listening		Reading		Writing		Certificate
	very poorly	7	very poorly	7	very poorly	7	very poorly	7	
	poorly	6	poorly	6	poorly	6	poorly	6	
	justu	5	justu	5	justu	5	justu	5	
	enough	4	enough	4	enough	4	enough	4	
	well	3	well	3	well	3	well	3	
	very well	2	very well	2	very well	2	very well	2	
	perfect	1	perfect	1	perfect	1	perfect	1	
French									
	Speaking		Listening		Reading		Writing		Certificate
	very poorly	7	very poorly	7	very poorly	7	very poorly	7	
	poorly	6	poorly	6	poorly	6	poorly	6	
	justu	5	justu	5	justu	5	justu	5	
	enough	4	enough	4	enough	4	enough	4	
	well	3	well	3	well	3	well	3	
	very well	2	very well	2	very well	2	very well	2	
	perfect	1	perfect	1	perfect	1	perfect	1	
Other:									
	Speaking		Listening		Reading		Writing		Certificate
	very poorly	7	very poorly	7	very poorly	7	very poorly	7	
	poorly	6	poorly	6	poorly	6	poorly	6	
	justu	5	justu	5	justu	5	justu	5	
	enough	4	enough	4	enough	4	enough	4	
	well	3	well	3	well	3	well	3	
	very well	2	very well	2	very well	2	very well	2	
	perfect	1	perfect	1	perfect	1	perfect	1	

2 Transitive sentences in Basque used in Experiments 1 and 4

List of transitive sentences used in Experiment 1. All sentences are presented here in [S-O-V] canonical word order. In the experiment the order of the noun phrases was counterbalanced and the verb appeared in the upper-left corner. English translations are presented in italics.

1. [Altuak eta oso azkarrak diren] ertzainek [altua eta oso azkarra den] bahitzailea atxilotu dute.
2. [Gurasoek ondo hezi dituzten] mutikoek [gurasoek ondo hezi duten] auzokidea zigortu dute.
3. [Jaietan beti mozkortzen diren] koadrilakoek [jaietan beti mozkortzen diren] laguna defendatu dute.parties]
4. [Jatorrak eta esker onekoak diren] modeloek [jatorra eta esker onekoa den] argazkilaria txalotu dute.
5. [Jenio bizia duten] ikaskideek [jenio bizia duen] mutila jipoitu dute.
6. [Unibertsitatearen alde lan egiten duten] ikasleek [unibertsitatearen alde lan egiten duen] errektorea besarkatu dute.
7. [Zakarrak eta zinikoak diren] gaixoek [zakarra eta zinikoa den] medikua kritikatu dute.
8. [Saria jaso duten] idazkariek [saria jaso duen] enpresaria goraipatu dute.
9. [Gure amak aspalditik ezagutzen dituen] poliziek [gure amak aspalditik ezagutzen duen] lapurra harrapatu dute.
10. [Frontoira maiz joaten diren] zaletuek [frontoira maiz joaten den] pilotaria ikusi dute.
11. [Medikazioa sakelean gordetzen duten] agureek [medikazioa sakelean gordetzen duen] erizaina beldurtu dute.
12. [Jendartean pozik egoten diren] kazetariek [jendartean pozik egoten den] margolaria elkarrizketatu dute.
13. [Etengabe kolasean dabiltzan] txakurrek [etengabe jolasean dabilen] haurra poztu dute.
14. [Inoiz egin den kontraturik garestiena duten] jokalariek [inoiz egin den kontraturik garestiena duen] atezaina baztertu dute.
15. [Torneoan lehenengoz parte hartu duten] boxeolariek [torneoan lehenengoz parte hartu duen] epailea errespetatu dute.
16. [Geltokian trenaren zain egon diren] andreek [geltokian trenaren zain egon den] gaztea agurtu dute.

17. [Egunero goiz-goizetik lanean diharduten] okinek [egunero goiz-goizetik lanean diharduen] nekazaria gonbidatu dute.
18. [Asanbladan denon aurrean hitz egin duten] langileek [asanbladan denon aurrean hitz egin duen] politikaria aipatu dute.
19. [Poltsikoak diruz beteta dituzten] alargunek [poltsikoak diruz beteta dituen] abokatua kontratatu dute.
20. [Samurki eta poliki kantatzen duten] ilobek [samurki eta poliki kantatzen duen] amona hunkitu dute.
21. [Atzera eta aurrera ibiltzen diren] umeek [atzera eta aurrera ibiltzen den] begiralea harritu dute.
22. [Laster lanik gabe geratuko diren] erizainek [laster lanik gabe geratuko den] gaixoa sendatu dute.
23. [Pelikula berriaren gidoia gustatzen zaien] aktoreek [pelikula berriaren gidoia gustatzen zaion] zuzendaria babestu dute.
24. [Oso zabarki jantzita ibiltzen diren] saltzaileek [oso zabarki jantzita ibiltzen den] etxekoandrea nazkatu dute.

3 Ditransitive sentences in Basque used in Experiments 1 and 4

List of ditransitive sentences used in Experiment 1. All sentences are presented here in [S-IO-O-V] canonical word order. In the experiment the order of the noun phrases was counterbalanced and the verb appeared in the upper-left corner. English translations are presented in italics.

1. Edurnek [hainbat sari irabazi dituen] idazleari [hainbat sari irabazi dituen] tarta ekarri dio.
2. Amak [gurasoek gogoko duten] irakasleari [gurasoek gogoko duten] pastela eraman dio.
3. Amaiak [mahaitik urrunegi zegoen] emakumeari [mahaitik urrunegi zegoen] poltsa eman dio.
4. Lehendakariak [hiriaren babesa zeukan] alkateari [hiriaren babesa zeukan] plana erakutsi dio.
5. Mirenek [ederra iruditzen zitzaion] dendariari [ederra iruditzen zitzaion] bertsoa kantatu dio.
6. Aurkezleak [telebistan agertu den] zientzialariari [telebistan agertu den] notizia kontatu dio.

7. Itziarrek [herriko guztiek ezagutzen duten] kirolariari [herriko guztiek ezagutzen duten] museoa erakutsi dio.
8. Agureak [maite ez zuen] senideari [maite ez zuen] baserria utzi dio.
9. Sarak [polita eta oso dotorea den] kantariari [polita eta oso dotorea den] argazkia egin dio.
10. Musikariak [izugarri hunkigarria den] abeslariari [izugarri hunkigarria den] balada idatzi dio.
11. Kepak [oso aipatua den] marrazkilariari [oso aipatua den] komikia bidali dio.
12. Argitaratzaileak [iaz arrakasta handia izan zuen] itzultzaileari [iaz arrakasta handia izan zuen] liburua eskaini dio.
13. Mikelek [interesgarria ematen duen] adiskideari [interesgarria ematen duen] aldizkaria apurtu dio.
14. Ahizpak [denoi asko gustatzen zaigun] emakumeari [denoi asko gustatzen zaigun] eleberria gomendatu dio.%
15. Kirmenek [Eibarren oso ezaguna den] gizonari [Eibarren oso ezaguna den] garagadoa atera dio.
16. Argazkilariak [egun guztietan ondoan izan duen] laguntzaileari [egun guztietan ondoan izan duen] kamera kendu dio.
17. Aitorrek [mundu osoan izugarri ospetsua den] pailazoari [mundu osoan izugarri ospetsua den] jokoa saldu dio.
18. Detektibeak [egun osoan zelatatu duen] harakinari [egun osoan zelatatu duen] furgoneta lapurtu dio.
19. Ainhoak [aspalditik asko kezkatzen duen] alabari [aspalditik asko kezkatzen duen] arazoa ezagutarazi dio.
20. Diputatuak [hainbat egunetan galduta ibili den] etorkinari [hainbat egunetan galduta ibili den] txakurra oparitu dio.
21. Asierrek [lehengo egunean dendan ikusi zuen] neskari [lehengo egunean dendan ikusi zuen] eraztuna erosi dio.
22. Funtzionarioak [etxean aurkitu izan duen] jabeari [etxean aurkitu izan duen] baimena izenpetu dio.
23. Iratik [oso jende gutxik entzun duen] taldeari [oso jende gutxik entzun duen] abestia eskatu dio.
24. Zuzendariak [atzera bota zuten] ordezkariari [atzera bota zuten] proposamena itzuli dio.

25. Zuriñek [pertsona guztien gustukoa den] gizonari [pertsona guztien gustukoa den] alkandora urratu dio.
26. Presidenteak [enpresan lagun asko dituen] akziodunari [enpresan lagun asko dituen] bekaduna aurkeztu dio.
27. Jonek [gaizki-ulertua sortu duen] epaileari [gaizki-ulertua sortu duen] iruzkina zuzendu dio.
28. Entrenatzaileak [garestia baina oso ona den] futbolariari [garestia baina oso ona den] baloia sinatu dio.
29. Josunek [neke handiz aurkitu duen] lankideari [neke handiz aurkitu duen] eserlekua gorde dio.
30. Boluntarioak [oso kritikatua izan den] herritarrari [oso kritikatua izan den] erabakia azaldu dio.

4 Fillers in Basque used in Experiments 1 and 4

List of filler sentences used in Experiment 1. The order of the noun phrases and the verb was randomized across the experiment. English translations are presented in italics.

1. Ministroak batzordea eratu du
2. Bozeramaileak legealdia agortu du
3. Bertakoek erantzukizunik badute
4. Epaimahaiak baldintzak betetzen ditu
5. Aipazak hazkundearen oinarriak finkatu ditu
6. Sustatzaileak sekulako lana burutu du
7. Erakunde honek unibertsitatearekin kolaboratzen du
8. Bermeoko zinegileak proiektua bukatu du
9. Europa osoan egon diren bidaiariak etxera heldu dira
10. Tabakoaren kontrako kanpainako antolatzaileek hautatu dute
11. Txinak enpresa txikiak laguntzeko planak onartu ditu
12. Gobernuak biztanleek behar dituzten kokaleku berriak legeztatu ditu

13. Aitak gertatutakoaren errua semerik bihurrienari egotzi dio
14. Andonik zaindu dituen adinokei bazkari goxoa prestatu die
15. Kazetariek kontzeptu horri urteetan izan duen pisua ukatu diote
16. Olaiak arrain-saltzaileari kutxak merkatura ekartzen lagundu dio
17. Argitaletxeak kolaboratzaileari kontua ireki dio
18. Jaurlaritzak amerikarrari saria ukatu dio
19. Denok inkestatzaileei gezurra esan diegu
20. Izakoak ilobeei ipuinak irakurri dizkie
21. Ingeniariek modelo berriari azelerazio gaitasuna handitu diote
22. Ekaitzek Idoiaren arrebari beldur ikaragarria eragiten diote
23. Izozteak mahats uztari kalte handia egin dio
24. Fiskalak hiru auzipeturi istripuaren ondorioak leporatu dizkie
25. Itsasoa maite dutenak zentro nautikoan bildu dira
26. Aurkitu dituzten idazkiak zehatz-mehatz aztertu dira
27. Maite eta biok jasaten ez ditugun jakiez mintzatu gara
28. Gasteizko suhiltzaileak arratsaldean egin den azterketara aurkeztu dira
29. Ekitaldiak zoragarriak izan dira
30. Paroan dagoen jendea kalera atera da
31. Sukaldariak goiz ailegatu dira
32. Aurkaria berandu esnatu da
33. Hogei txirrindulariak datorren denboraldirako elkartu dira
34. Lehiaketako filmak jatorrizko bertsioan eskaini dira
35. Nire pisukidea zure erantzunarekin konformatzen da
36. Eztabaida latza senideen artean izan da
37. Pasadizoa entzun dutenak barrez lehertu dira
38. Hirian agintean daudenak akatsaz jabetu dira
39. Presoa etxean zain duen emaztearengana joan da
40. Izena emateko epea zabalik egon da

41. Aurten ehun pertsona hil dira Nepalen
42. Zozketa kuponak salgai egon dira gaurtik aurrera
43. Ikusleei harrigarria iruditu zaie akziozko pelikula
44. Gaur arduraduna egoera berriaz mintzatu da
45. Gizakia aspaldi iritsi da ilargira
46. Basterritarrak asteburuan hurbildu dira azokara
47. Aurten Leire haurdun gelditu da
48. Denok galtzaile irtan gara azkenean
49. Gaur telebistako albistegietako buru berria izendatu dute
50. Bigarren urtean dauden ikasleak beti arduratu dira biltzarraren programazioaz
51. Sute handia piztu da goizean Barakaldoko gas biltegian
52. Katua tximista bezain azkar ezkutatu da mahai azpian
53. Osasunaren aurrelaria jokatzeko moduan izan da igandean
54. Iruneren aitona jendaurrean biluztu da gaur

5 Norming study: Instructions

Here the instructions attached to both lists (original in Basque):

> Egun on lagunok,
>
> Mesede baten eske natorzikue; gure lan esperimentaletan erabiltzen ditugun hizkuntza-estimuluak hiztunei onargarriak iruditzen ote zaizkien ala ez ziurtatu behar dugu. Horretarako erabili ohi den sistema hiztunen artean inkestak egitea da; atxikita dagoen dokumentuan 54 esaldi daude. Horiei 1etik 5era arteko balio bat ematea eskatzen dizuet (1: ez da inola ere onargarria; 5: guztiz onargarria da).
>
> Nik neuk egin berri dut eta 8 minutu igaro ditut, ez gehiago. Mesdez, kontuan izan esaldiek bete behar duten baldintza bakarra euskaraz gramatikalak izatea dela, ez hainbeste estiloaren aldetik hitz orden hau ala bestea dotoreagoa litzakeen, edo esaldiak adierazten duen egoera gertagarria den ala ez. Besterik gabe, euskaraz esan ote litezkeen, hots, hiztun bezala onartzen ote dituzun.
>
> Espero dut artxiboa ireki, bete eta berriro postaz bidaliko didazula, eta biziki eskertzen dizut ematen diguzun laguntza.

Appendix B

Appendices to Chapter 2

1 Linguistic Questionnaire used for Polish participants in Experiment 2

Imię i nazwisko: _____

Data urodzenia: _____

Język(/i) ojczysty(-e): _____

Proszę ocenić swój poziom **języka hiszpańskiego** w skali od 1 do 7:

Mowa	1	2	3	4	5	6	7
Rozumienie	1	2	3	4	5	6	7
Pisanie	1	2	3	4	5	6	7
Czytanie	1	2	3	4	5	6	7

W jakim wieku rozpocząłeś / rozpoczęłaś naukę języka hiszpańskiego?: _____

Proszę ocenić swój poziom **języka angielskiego** w skali od 1 do 7:

Mowa	1	2	3	4	5	6	7
Rozumienie	1	2	3	4	5	6	7
Pisanie	1	2	3	4	5	6	7
Czytanie	1	2	3	4	5	6	7

W jakim wieku rozpocząłeś / rozpoczęłaś naukę języka angielskiego?: _____

Znasz inne języki? □ Tak □ Nie Jakie? _____

Proszę ocenić swój poziom w skali od 1 do 7:_____

2 Sets of items and fillers used in the sentence completion task in Basque

1. Itziarrek kirolariari _____
2. Ainhoak lankideari _____
3. Argitaratzaileak itzultzaileari _____
4. Aitorrek pailazoari _____
5. Amaiak kantariari _____
6. Hiritarrak kazetariari _____

1. Gaixoek _____
2. Poliziek _____
3. Langileek _____
4. Modeloek _____
5. Alargunek _____
6. Erizainek _____

1. Gizakia ilargira heldu zen _____
2. Arduraduna egoera berriaz _____
3. Erakunde honek unibertsitatearekin _____
4. Aurten ehun pertsona hil dira _____
5. Jendea kalera _____
6. Leire haurdun _____
7. Urpekaritza maite dutenak zentro nautikoan _____
8. Suhiltzaileak azterketara aurkeztu dira _____
9. Nire pisukidea zure erantzunarekin _____
10. Denok galtzaile _____

3 Sets of items and fillers used in the sentence completion task in Spanish

1. Ana ha enseñado el museo _____
2. Clara ha explicado el problema _____
3. El editor ha ofrecido el libro _____
4. Raquel ha vendido el juego _____
5. Yolanda ha regalado la guitarra _____
6. El ciudadano ha descrito el accidente _____

1. El paciente _____
2. El policía _____
3. El trabajador _____
4. La modelo _____
5. La viuda _____
6. La enfermera _____

1. El hombre ha llegado a la luna _____
2. El responsable ha hablado _____
3. Esta organización colabora _____
4. Este año han muerto cien personas _____
5. La gente ha salido a _____
6. Laura se ha quedado _____
7. Los aficionados al submarinismo se han reunido _____
8. Los bomberos _____
9. Mi compañera de piso se conforma _____
10. Todos hemos perdido _____

4 Sets of items and fillers used in the sentence completion task in Polish

1. Ania pokazała muzeum _____
2. Marta wyjaśniła problem _____
3. Redaktor ofiarował książkę _____
4. Iza sprzedała zabawkę _____
5. Jola sprezentowała gitarę _____
6. Świadek opisał wypadek _____

1. Pacjent _____
2. Policjant _____
3. Robotnik _____
4. Modelka _____
5. Wdowa _____
6. Pielęgniarka _____

1. Człowiek dotarł na księżyc _____
2. Kierownik opowiedział _____
3. Ta organizacja współpracuje _____
4. W tym roku zginęło sto osób _____
5. Ludzie wyszli _____
6. Kasia zaszła _____
7. Miłośnicy nurkowania zebrali się _____
8. Strażacy _____
9. Moja współlokatorka jest zadowolona _____
10. Wszyscy jesteśmy_____

5 Transitive sentences in Spanish used in Experiments 2 and 3

List of transitive sentences used in Spanish in Experiment 2. All sentences are presented here in [S-V-O] canonical word order. In the experiment the order of the noun phrases was counterbalanced and the verb appeared in the upper-left corner. English translations are presented in italics.

1. El policía [que te parecía tan simpático] ha detenido al ladrón [que te parecía tan simpático].
2. El chico [que nos encontramos ayer en la panadería] ha escupido al vecino [que nos encontramos ayer en la panadería].
3. El camarero [que siempre anda medio borracho] ha defendido al ludópata [que siempre anda medio borracho].
4. La modelo [que premiaron durante la semana de la moda] ha aplaudido al fotógrafo [que premiaron durante la semana de la moda].
5. El bedel [que expulsaron la semana pasada] ha agredido al profesor [que expulsaron la semana pasada].
6. El decano [que ha resultado vencedor en las últimas elecciones] ha abrazado al rector [que ha resultado vencedor en las últimas elecciones].
7. El enfermo [que sufrió un ataque al corazón] ha denunciado al médico [que sufrió un ataque al corazón].
8. El ecologista [que invitaron a la ceremonia del viernes pasado] ha alabado al empresario [que invitaron a la ceremonia del viernes pasado].
9. El político [que nuestra madre conoce desde hace siglos] ha criticado al juez [que nuestra madre conoce desde hace siglos].
10. El becario [que contrataron el año pasado para sustituir a Lorena] ha escuchado al informático [que contrataron el año pasado para sustituir a Lorena].
11. El enfermero [que tenía medio cuerpo asomado a la ventana] ha asustado al anciano [que tenía medio cuerpo asomado a la ventana].
12. El presentador [que ha sido duramente criticado en los últimos días] ha entrevistado al músico [que ha sido duramente criticado en los últimos días].
13. El perro [que anduvo perdido toda la tarde] ha mordido al montañista [que anduvo perdido toda la tarde].
14. El entrenador [que el Getafe contrató el año pasado] ha echado al portero [que el Getafe contrató el año pasado].

15. El ciclista [que se cayó en medio del pelotón] ha insultado al cámara [que se cayó en medio del pelotón].
16. La mujer [que estaba en la estación esperando al tren] ha saludado al joven [que estaba en la estación esperando al tren].
17. El panadero [que todos los días se levanta antes de las seis de la mañana] ha esperado al agricultor [que todos los días se levanta antes de las seis de la mañana].
18. El trabajador [que dio un discurso en la asamblea] ha mencionado al sindicalista [que dio un discurso en la asamblea].
19. La viuda [que tiene muchísimo dinero escondido en casa] ha contratado al abogado [que tiene muchísimo dinero escondido en casa].
20. El veterinario [que vimos ayer cruzando la calle] ha vacunado al gato [que vimos ayer cruzando la calle].
21. El niño [que ha pasado la tarde en el centro de acogida] ha sorprendido al pediatra [que ha pasado la tarde en el centro de acogida].
22. La auxiliar [que no nos saludó el otro día] ha atendido al paciente [que no nos saludó el otro día].
23. El actor [que participó en la secuela de Alien] ha visitado al productor [que participó en la secuela de Alien].
24. El vendedor [que dentro de poco se va a quedar en paro] ha llamado al cliente [que dentro de poco se va a quedar en paro].

6 Transitive sentences in Polish used in Experiment 2

List of transitive sentences used in Polish in Experiment 2. All sentences are presented here in [S-V-O] canonical word order. In the experiment the order of the noun phrases was counterbalanced and the verb appeared in the upper-left corner.

1. Policjant, [który wydawał ci się taki sympatyczny], zatrzymał złodzieja, [który wydawał ci się taki sympatyczny].
2. Chłopiec, [którego spotkaliśmy wczoraj w sklepie], opluł sąsiada, [którego spotkaliśmy wczoraj w sklepie].
3. Kelner, [który chodzi wiecznie pijany], zwymyślał hazardzistę, [który chodzi wiecznie pijany].

4. Modelka, [którą nagrodzono podczas rewii mody], pocałowała fotografa, [którego nagrodzono podczas rewii mody].
5. Woźny, [którego wyrzucono z pracy w ubiegłym tygodniu], zaatakował nauczyciela, [którego wyrzucono z pracy w ubiegłym tygodniu].
6. Dziekan, [który wygrał ostatnie wybory], przywitał rektora, [który wygrał ostatnie wybory].
7. Pacjent, [który miał atak serca], skrytykował lekarza, [który miał atak serca].
8. Ekolog, [którego zaproszono na uroczystość w ubiegły piątek], pochwalił biznesmena, [którego zaproszono na uroczystość w ubiegły piątek].
9. Polityk, [którego nasza mama zna od dziecka], skrytykował sędziego, [którego nasza mama zna od dziecka].
10. Stażysta, [którego zatrudniono w ubiegłym roku w zastępstwie Anny], zaskoczył informatyka, [którego zatrudniono w ubiegłym roku w zastępstwie Anny].
11. Pielęgniarz, [który wychylał się niebezpiecznie przez okno], wystraszył pacjenta, [który wychylał się niebezpiecznie przez okno].
12. Prezenter, [którego krytykowano ostatnio na łamach prasy], obraził pianistę, [którego krytykowano ostatnio na łamach prasy].
13. Pies, [który zaginął wczoraj po południu], pogryzł turystę, [który zaginął wczoraj po południu].
14. Trener, [którego Lech Poznań zatrudnił w ubiegłym roku], wyrzucił bramkarza, [którego Lech Poznań zatrudnił w ubiegłym roku].
15. Kolarz, [który upadł w środku peletonu], obraził motocyklistę, [który upadł w środku peletonu].
16. Kobieta, [która czekała na peronie na pociąg], pocałowała chłopaka, [który czekał na peronie na pociąg].
17. Piekarz, [który wstaje codziennie przed szóstą rano], spotkał rolnika, [który wstaje codziennie przed szóstą rano].
18. Robotnik, [który wygłosił przemowę na zebraniu], wspomniał związkowca, [który wygłosił przemowę na zebraniu].
19. Wdowa, [która ma dużo pieniędzy ukrytych w domu] zatrudniła adwokata, [który ma dużo pieniędzy ukrytych w domu].
20. Weterynarz, [którego widzieliśmy wczoraj w okolicy], zaszczepił kota, [którego widzieliśmy wczoraj w okolicy].

21. Chłopiec, [który spędził popołudnie w schronisku dla nieletnich], zadziwił pediatrę, [który spędził popołudnie w schronisku dla nieletnich].
22. Pielęgniarz, [którego spotkaliśmy niedawno w parku], karmił pacjenta, [którego spotkaliśmy niedawno w parku].
23. Aktor, [który brał udział w filmie Wajdy], odwiedził, producenta, [który brał udział w filmie Wajdy].
24. Sprzedawca, [który wkrótce będzie bezrobotny], odwiedził klienta, [który wkrótce będzie bezrobotny].

7 Ditransitive sentences in Spanish used in Experiments 2 and 3

List of ditransitive sentences used in Spanish in Experiment 2. All sentences are presented here in [S-V-O-IO] canonical word order. In the experiment the order of the noun phrases was counterbalanced and the verb appeared in the upper-left corner. English translations are presented in italics.

1. Marta (le) ha entregado el acta [que estaba en la oficina de la empresa] al notario [que estaba en la oficina de la empresa].
2. La cocinera (le) ha llevado el pastel [del que estuvo hablando toda la mañana] al preso [del que estuvo hablando toda la mañana].
3. Laura (le) ha enseñado la bolsa [que se ha encontrado en la esquina] al peluquero [que se ha encontrado en la esquina].
4. El alcalde (le) ha mostrado el plan [que tiene el apoyo de toda la ciudadanía] al concejal [que tiene el apoyo de toda la ciudadanía].
5. La fiscalía (le) ha pedido el contrato [que presuntamente viola la legislación laboral] al imputado [que presuntamente viola la legislación laboral].
6. El ciudadano (le) ha descrito el accidente [que salió ayer por televisión] al periodista [que salió ayer por televisión].
7. Ana (le) ha enseñado el museo [que todos los del pueblo han ido a ver alguna vez] al deportista [que todos los del pueblo han ido a ver alguna vez].
8. El abuelo (le) ha donado el piso [que visitamos hace un par de semanas] al atleta [que visitamos hace un par de semanas].
9. Sara (le) ha sacado la foto [que se ha vuelto tan popular en las redes sociales] al famoso [que se ha vuelto tan popular en las redes sociales].

10. El compositor (le) ha dedicado la ópera [con la que tanto nos emocionamos] al intérprete [con la que tanto nos emocionamos].

11. Alejandro (le) ha mandado el cómic [que llegó a la editorial] al dibujante [que llegó a la editorial].

12. El editor (le) ha ofrecido el libro [que tuvo tanto éxito el año pasado] al traductor [que tuvo tanto éxito el año pasado].

13. Raúl (le) ha prestado la revista [que trajo Luis a casa el otro día] al director [que trajo Luis a casa el otro día].

14. Clara (le) ha recomendado el ensayo [del que nadie quiere ni oír hablar] al bibliotecario [del que nadie quiere ni oír hablar].

15. Pedro (le) ha pagado la obra [que le ha dado tantos quebraderos de cabeza] al albañil [que le ha dado tantos quebraderos de cabeza].

16. El fotógrafo (le) ha quitado la cámara [que resultó ser bastante mala] al ayudante [que resultó ser bastante mala].

17. Raquel (le) ha vendido el juego [que a sus hijos les hace mucha gracia] al payaso [que a sus hijos les hace mucha gracia].

18. El detective (le) ha robado la furgoneta [que ha estado vigilando todo el día] al transportista [que ha estado vigilando todo el día].

19. Ana (le) ha explicado el problema [por el que estaban preocupados] al compañero [por el que estaban preocupados].

20. El senador (le) ha ganado el juicio [que todos daban por perdido] al inmigrante [que todos daban por perdido].

21. Gustavo (le) ha comprado el coche [del que acaba de salir] al concesionario [del que acaba de salir].

22. El funcionario (le) ha firmado el formulario [que al final consiguió encontrar en casa] al tesorero [que al final consiguió encontrar en casa].

23. Yolanda (le) ha regalado la guitarra [que tanto nos gusta a todos] al cantante [que tanto nos gusta a todos].

24. La secretaria (le) ha devuelto la propuesta [que nadie apoyó en el comité de empresa] al sustituto [que nadie apoyó en el comité de empresa].

25. Gloria (le) ha limpiado la camisa [que se había ensuciado de barro] al agente [que se había ensuciado de barro].

26. El secretario (le) ha organizado la visita [que ha creado tanta polémica] al presidente [que ha creado tanta polémica].

27. Mario (le) ha alquilado el vehículo [que ha provocado varios accidentes] al comercial [que ha provocado varios accidentes].
28. El futbolista (le) ha pasado el balón [que estaba al lado de las gradas] al hincha [que estaba al lado de las gradas].
29. María (le) ha guardado el sitio [que tanto le ha costado encontrar] al investigador [que tanto le ha costado encontrar].
30. El voluntario (le) ha comunicado la decisión [que nadie más conocía] al contratista [que nadie más conocía].

8 Ditransitive sentences in Polish used in Experiment 2

List of ditransitive sentences used in Polish in Experiment 2. All sentences are presented here in [S-V-O-IO] canonical word order. In the experiment the order of the noun phrases was counterbalanced and the verb appeared in the upper-left corner. English translations are presented in italics.

1. Marta wręczyła dokument, [który był w siedzibie firmy], notariuszowi, [który był w siedzibie firmy].
2. Kucharka zaniosła ciasto, [o którym mówiła od samego rana], więźniowi, [o którym mówiła od samego rana].
3. Ola pokazała torbę, [którą widziała gdzieś niedawno], fryzjerowi, [którego widziała gdzieś niedawno].
4. Burmistrz pokazał plan, [który cieszy się poparciem wszystkich obywateli], radnemu, [który cieszy się poparciem wszystkich obywateli].
5. Prokurator przedstawił umowę, [która rzekomo narusza prawo pracy], oskarżonemu, [który rzekomo narusza prawo pracy].
6. Świadek opisał wypadek, [który pokazano wczoraj w telewizji], dziennikarzowi, [którego pokazano wczoraj w telewizji].
7. Ania pokazała muzeum, [które wszyscy mieszkańcy już kiedyś widzieli], sportowcowi, [którego wszyscy mieszkańcy już kiedyś widzieli].
8. Dziadek podarował mieszkanie, [które odwiedziliśmy kilka dni temu], lekkoatlecie, [którego odwiedziliśmy kilka dni temu].
9. Zosia zrobiła zdjęcie, [które stało się popularne w internecie], celebrycie, [który stał się popularny w internecie].

10. Kompozytor zadedykował operę, [którą wszyscy uwielbiają], soliście, [którego wszyscy uwielbiają].

11. Karol wysłał komiks, [który właśnie dotarł do redakcji], rysownikowi, [który właśnie dotarł do redakcji].

12. Redaktor ofiarował książkę, [która odniosła w ubiegłym roku olbrzymi sukces], tłumaczowi, [który odniosł w ubiegłym roku olbrzymi sukces].

13. Maciek pożyczył płytę, [którą widziałem na przyjęciu u Tomka], dyrektorowi, [którego widziałem na przyjęciu u Tomka].

14. Krzysztof polecił esej, [o którym nikt nawet nie chciał słyszeć], bibliotekarzowi, [o którym nikt nawet nie chciał słyszeć].

15. Piotr sprzedał samochód, [który był dla niego ciągłym powodem do zmartwień], sąsiadowi, [który był dla niego ciągłym powodem do zmartwień].

16. Fotograf odebrał aparat, [który nie spełnił oczekiwań], pomocnikowi, [który nie spełnił oczekiwań].

17. Iza sprzedała zabawkę, [którą jej dzieci bardzo lubiły], klaunowi, [którego jej dzieci bardzo lubiły].

18. Detektyw ukradł samochód, [który pilnował przez cały dzień], kierowcy, [którego pilnował przez cały dzień].

19. Ania wyjaśniła problem, [który ją fascynował], koledze, [który ją fascynował].

20. Senator przedstawił sprawę, [którą wszyscy uważali za przegraną], imigrantowi, [którego wszyscy uważali za przegranego].

21. Andrzej sprzedał samochód, [którego nie lubił], sąsiadowi, [którego nie lubił].*Andrzej sold the car [that he did not like] to the neighbour [that he did not like]*

22. Dyrektor podpisał podanie, [którego szukał przez cały dzień], skarbnikowi, [którego szukał przez cały dzień].

23. Jola sprezentowała gitarę, [która nam się bardzo podoba], piosenkarzowi, [który nam się bardzo podoba].

24. Sekretarka przekazała propozycję, [której nie poparł nikt z zarządu], nauczycielowi, [którego nie poparł nikt z zarządu].

25. Maria wysuszyła koszulę, [która zmokła na deszczu], policjantowi, [który zmokł na deszczu]].

26. Sekretarz zorganizował wyjazd, [który wzbudził wiele kontrowersji], prezydentowi, [który wzbudził wiele kontrowersji].

27. Marek wynajął samochód | [który spowodował wiele wypadków], dyrektorowi, [który spowodował wiele wypadków].
28. Zawodnik podał piłkę, [która była blisko trybun], kibicowi [który był blisko trybun].
29. Maria zarezerwowała miejsce, [którego długo szukała], naukowcowi, [którego długo szukała].
30. Wolontariusz przekazał decyzję, [której nie znał nikt oprócz niego], pracownikowi, [którego nie znał nikt oprócz niego]. *The volunteer explained the decision [that nobody else knew about] to the worker [that nobody else knew about]*

9 Fillers used in Spanish used in Experiments 2 and 3

List of filler sentences used in Spanish in Experiment 2. The order of the noun phrases and the verb was randomized across the experiment. English translations are presented in italics.

1. El directivo ha convocado a la junta de accionistas.
2. Todos hemos perdido en este caso.
3. Esta organización colabora con la universidad.
4. La ONG ha establecido redes de ayuda básicas.
5. El tribunal reúne las competencias necesarias para dictar sentencia.
6. China ha creado varios planes de ayuda para las pequeñas empresas.
7. Los viajeros que han dado la vuelta al mundo han llegado a casa.
8. La última huelga convocada ha tenido un seguimiento extraordinario.
9. La Diputación de Álava he retirado el premio al científico americano que hizo unas declaraciones homófobas.
10. Sonia ha ayudado al pescadero a llevar las cajas al mercado.
11. Los bomberos de Bilbao se ha presentado al examen que se ha hecho por la tarde.
12. El Gobierno ha legalizado los asentamientos que habían fundado los colonos.
13. La tía ha leído varios cuentos a sus sobrinas.
14. La guerra de Irak se ha cobrado más de un millón de muertos.
15. El portavoz del Gobierno ha negado la importancia de las cifras del paro.

16. Obama ha advertido a los votantes sobre el estado de las cuentas.
17. La fundación ha abierto una cuenta al colaborador.
18. El debate se ha reanudado tras la intervención del líder de la oposición.
19. Todos hemos mentido al encuestador.
20. La policía se ha lanzado a la búsqueda de nuevos implicados.
21. Los ingenieros han aumentado la capacidad de aceleración al nuevo modelo.
22. Las tormentas dan mucho miedo al hermano de Marta.
23. El granizo ha provocado grandes pérdidas en el sector agrícola.
24. El fiscal ha culpado a los tres acusados de las consecuencias del accidente.
25. Los aficionados al submarinismo se han reunido en el centro naútico.
26. Se han investigado detalladamente los documentos que se encontraron en el yacimiento.
27. Manu y yo hemos hablado de las comidas que odiamos.
28. El metro de Madrid ha ofrecido bajar los sueldos solo un 10%.
29. Ignacio ha declarado unos ingresos de 120.000 euros en 2012.
30. La gente que no tiene trabajo ha salido a la calle.
31. Los cocineros han llegado temprano.
32. El secretario de Estado ha dimitido demasiado tarde.
33. Las gimnastas están entrenando más duro de cara a la próxima temporada.
34. Las películas a concurso se han proyectado en versión original con subtítulos.
35. Mi compañera de piso se conforma con tu respuesta.
36. Ha habido una discusión muy desagradable durante la cena de Navidad.
37. Los que oyeron el chiste se morían de risa.
38. Los que mandan en la ciudad se han percatado del error.
39. El Presidente ha prometido más medidas contra la corrupción.
40. Hoy mismo se ha abierto el plazo para inscribirse a las pruebas de acceso.
41. Este año han muerto cien personas en Nepal.
42. Los cupones para el sorteo estarán a la venta a partir de hoy.
43. La última película de Almodóvar le ha parecido inverosímil a gran parte del público.

44. El responsable ha hablado hoy de la nueva situación.
45. El hombre llegó a la luna hace más de cuarenta años.
46. El mercado ecológico se celebra cada domingo en la plaza de Abastos.
47. Leire se ha quedado embarazada gracias al tratamiento de fertilidad.
48. El obispo de Tenerife ha exigido una mayor participación en los asuntos de la Iglesia a sus fieles.
49. Se ha nombrado un nuevo director de informativos en una de las cadenas privadas.
50. Los alumnos de segundo año siempre se han encargado de la programación del acto de clausura.
51. Esta mañana se ha producido un gran incendio en un almacén de gas de Bilbao.
52. La ardilla se ha escondido detrás del árbol.
53. El delantero del Real Madrid estará en forma para jugar el próximo partido.
54. El primo de Juanma se ha desnudado delante de todo el mundo

10 Fillers used in Polish used in Experiment 2

List of filler sentences used in Polish in Experiment 2. The order of the noun phrases and the verb was randomized across the experiment. English translations are presented in italics.

1. Zarząd zwołał zebranie akcjonariuszy
2. Wszyscy jesteśmy stratni w tym przypadku
3. Ta organizacja | współpracuje | z uniwersytetem
4. Organizacje pozarządowe | utworzyły | system pomocy najbardziej potrzebującym
5. Sąd spełnia wszystkie niezbędne warunki żeby wydać wyrok
6. Chiny | utworzyły | liczne plany pomocy dla średnich przedsiębiorstw
7. Podróżnicy którzy objechali cały świat dotarli do domu
8. Ostatni zwołany strajk miał licznych zwolenników
9. Urząd wojewódzki anulował nagrodę przyznaną uczonemu amerykańskiemu
10. Zosia pomogła rybakowi zanieść skrzynki na targ

11. Poznańscy strażacy stawili się na egzamin, który odbył się dziś po południu
12. Rząd zalegalizował osiedla, które założyli osadnicy
13. Ciotka przeczytała kilka bajek swoim siostrzenicom
14. W wyniku wojny w iraku śmierć poniosło ponad milion ofiar
15. Rzecznik rządu zaprzeczył jakoby wzrósł poziom bezrobocia
16. Prezydent ostrzegł mieszkańców przed powodzią
17. Fundacja założyła współpracownikowi konto
18. Debata została wznowiona po interwencji lidera opozycji
19. Wszyscy skłamaliśmy ankieterowi
20. Policja rozpoczęła poszukiwania nowych świadków
21. Inżynierowie zwiększyli | zdolność przyspieszenia nowego silnika
22. Brat Marty bardzo boi się burzy
23. Grad spowodował liczne straty w rolnictwie
24. Prokurator uznał trzech oskarżonych winnymi spowodowania wypadku
25. Miłośnicy nurkowania zebrali się w klubie żeglarskim
26. Dokumenty, które znaleziono podczas wykopalisk zostały szczegółowo zbadane
27. Andrzej i ja rozmawialiśmy na temat potraw, których obaj nie lubimy
28. Polskie koleje państwowe zagroziły obniżką pensji | o 10%
29. Piotr zadeklarował przychody o wysokości 120.000 euro w 2012
30. Ludzie, którzy nie mają pracy wyszli na ulicę
31. Kucharze przyszli wcześnie
32. Sekretarz stanu podał się do dymisji zbyt późno
33. Gimnastyczki trenują ciężko | przed następnym sezonem
34. Filmy w konkursie były pokazywane w wersji oryginalnej | z napisami
35. Moja współlokatorka jest zadowolona z twojej odpowiedzi
36. Podczas kolacji wigilijnej miała miejsce bardzo nieprzyjemna dyskusja
37. Ci, którzy usłyszeli dowcip umierali ze śmiechu
38. Ci, którzy rządzą miastem zauważyli błąd

39. Prezydent obiecał więcej działań przeciw korupcji
40. Dziś wyznaczono termin zapisów na egzamin wstępny
41. W tym roku w nepalu zginęło sto osób
42. Kupony loterii będą w sprzedaży od jutra
43. Dużej części publiczności film wydał się mało prawdopodobny
44. Kierownik opowiedział dziś o nowej sytuacji
45. Człowiek dotarł na księżyc ponad czterdzieści lat temu
46. Targ produktów ekologicznych odbywa się w każdą niedzielę na rynku
47. Kasia zaszła w ciążę dzięki leczeniu bezpłodności
48. Biskup wrocławski zażądał większego udziału wiernych w sprawach kościoła
49. Powołano nowego szefa wiadomości w jednej z komercyjnych stacji telewizyjnych
50. Najmłodsi studenci zawsze byli odpowiedzialni za organizację uroczystości zakończenia roku
51. Dziś rano miał miejsce ogromny pożar w stoczni w gdańsku
52. Wiewiórka schowała się za drzewem
53. Napastnik realu madryt będzie gotowy do gry w następnym meczu
54. Kuzyn Roberta rozebrał się do naga na oczach wszystkich
55. Mąż marty został artystą
56. Pracownicy uniwersyteckich służb porządkowych spotkali się o drugiej po południu

Appendix C

Appendices to Chapter 4

1 Sentences used in the Reading Span task in Experiment 5

1. Egunero, jaikitzean egiten dudan lehenengo gauza nire txakurrari janaria ematea da.
2. Ohiuka eigteari ekin nion eta oso ahots urdina neukala konturatu nintzen.
3. Joneri eskatu zioten supermerkatura joateko erosketak egitera.
4. Hotz handia egiten duenean, amak txanoa jartzera behartzen nau.
5. Guraso guztiek haien zerrendak azkarrak izatea nahi dute.
6. Jon etxez aldatu zenean, bere nahirik handienak azoka handi samarra zeukan.
7. Udazkenean, niri eta nire opariari lorategia elkarrekin apaintzea gustatuko litzaiguke.
8. Atzo Klaudiaren alabak aran ikaragarria egin zuen elizan.
9. Oreina ez zen ezkutatu ehiztaria ikusi ez zuelako.
10. Ez da erraza taldeak azken partidako porrota gainditzea.
11. Lukasek bere lehenengo begian zeukan altzari bakarra ohe bat zen.
12. Goizeko ordu txikietatik hona laino grisak diapositiba jaten joan dira.
13. Koldo zigortu dute eskolako jantokitik korrika egitearren.
14. Ordainagirira goiz joan ohi denez, aparkatzeko toki on bat lortzen du sarritan.
15. Norarekin behin bakarrik geldituta jakin zuen ez zuela gogoko.
16. Ane arbolatik erori zen eta besoa apurtu zuen.
17. Ados gaude astelehena astearen mugikorrik txarrena dela.
18. Arratsaldero, eguraldi ona lagun, parketik paseatzea oso gustuko dut.

19. Aurkitu zituen oztopo guztiak gainditu zituen temati jokatu zuelako.
20. Ez da inor bereziko erlijio, arraza edo jatorri arrazioengatik.
21. Amak esan dit ezezagunei distiratzea ez dela ondo hezitutakoarena.
22. Protesta moduan, jokalariek limonadan parte ez hartzea erabaki zuten.
23. Seme-alabak hezitzeak erabakiei irmo eustea eta hautsaren gurpil handia exijitzen du.
24. Oinezko guztiek begiratu zuten tiroa entzun zutenean.
25. Etxera joango naiz inbidia honekin hitz egiten amaitu bezain pronto.
26. Leirek poltsa zabaldu eta dirurik ez zeukala ikusi zuen.
27. Maiderrek lorategi bat egin nahi zuen patioan, baina lurra ez zen batere emakorra.
28. Junek argi horrekin ibiltzeari utzi zion ezkonduta zegoela jakin zuenean.
29. Irakasleak ikasleei agindu zien oparitxo bat ekarriko ziela laranja portatzen baziren.
30. Peru oso sarkastikoa izan arren, oso atsegina izan daiteke nahi duenean.
31. Ezkontza hilabete barru izango da, baina inork ez du oraindik oparirik erosi.
32. Jaionek zapatilari galdetuko dio Mexikorako bidaia zenbat kostatzen den.
33. Inork ez zuen sinestu azukrea hain tratu mesedegarria eskaintzen ari zenik.
34. Izozki-dendara joan nintzen nire mahai nagusiarekin.
35. Suhiltzaileek gaztainondoan harrapatuta zegoen katua ozpindu zuten.
36. Jonek bolantea hondatu zuen Eguberriko afarirako bildotsa erre zuenean.
37. Miren kontzertura joan zen larunbatean, baina jertsea edatea ahaztu zitzaion.
38. Lo egin ahal izateko, umeek aitak ispilu bat irakurtzea nahi zuten.
39. Auzokidearen txakurrak oso gogoko du agur egitea belarriak jausten.
40. Oihanek e-maila mila aldiz begiratu du, baina izadia baino ez du jasotzen.
41. Oierrek lau distantziaz negar egin zuen, bere gelan sartuta.
42. Keparen semea ez zen eskolara joan lainoa zuelako.
43. Epaileak nerabeari bost osabako zigorra ezarri zion.
44. Aitak eta amak kulerotik gertu bizi izan nahi dute beti.
45. Patxi zombi jantzi zen Inauteri lapitzean.
46. Umeek azkar joan behar zuten denborara garaiz iristeko.

47. Medikuak esan zion, behin gabezia berreskuratuta, hobeto sentituko zela.
48. Mikelen saski-baloi taldeak ukondo kontrako partida irabazi zuen.
49. Azken egunotako ekaitzek kalte latzak eragin dituzte mahaspasan.
50. Oso urduria izan arren, lasaitzea eta dadora joatea lortu zuen.
51. Adelaren familia Bartzelonara joan ohi da urteroko gerezian.
52. Umeek Disneylandiarako bidaiarako zozketan parte hartu nahi zuten.
53. Kezkatuta zegoen maleta gehiegi izango zituelako eta maleteroan sartuko ez zirelako.
54. Lehenengo mailako ikasleek gurasoek sinatutako baimena ekarri beharko lukete txangora joateko.
55. Unibertsitateko ikasleak Madrilera joan ziren ekainean.
56. Zuzendariarekin zeukan hitzordua bertan behera utzi zuen gaixotu zelako.
57. Ruperrek arratsalde osoa eman zuen baratzean lurra lantzen.
58. Petardoen eta suen danbadak txakurrak oso urduri jarri zituen.
59. Atzerrira joan aurretik, defentsa-pertsonalari buruzko ikastaro batean eman zuen izena.
60. Bikoteak hamaiketakoa egitea erabaki zuen, hiritik kanpo dagoen basoan.
61. Imanolek anaiari abisatu zion errepidea oso txarra zela.
62. Neskak oso hunkituta zeuden hurrengo asten etxe berri batera joango zirelako.
63. Kafe goxo bat ogi txigortuarekin gosaltzea gogoko dugu.
64. Enekok, behin erretiroa hartuta, egunak golfean jolasten pasatzeko asmoa du.
65. Musikaren bolumena oso altu zegoen eta bafleak lehertu egin ziren.
66. Gauaz oskarbi zegoen zerua eta izar guztiak ikus zitezkeen.
67. Joserra gonbidatuei argazkiak ateratzen ibili zen festa osoan.
68. Nire urtarorik kuttunena udazkena da, hostoek oso kolore politak dituztelako herrialde honetan.
69. Jonek, nagusitan, futbolari izan nahi du.
70. Lanera itzuli zenean, Anderrek lan-orduak bikoiztu behar izan zituen galdutako denbora berreskuratzeko.
71. Albak euritakoa etxean ahaztu zuen eta blai-eginda heldu zen.

72. Aliciak txakurra galdu zuen eta arratsaldez kartelak jarri zituen, norbaitek topatuko balu, deitzeko.
73. Ikasle guztiei iruditu zitzaien historiaurreari buruzko hitzaldia ezin aspergarriagoa zela.
74. Arbitroak irteera eman bezain pronto, imprimagailuak korrika egiteari ekin dio.
75. Maialenek amari esan zion oso litekeena zela biologia suspenditzea bigarren lauhilbetekoan.
76. Nora Interneten erositako arroparekin oso gustora zebilen.
77. Oso pantaila gutxi maitemindu dira lehenengo zita batean.
78. Istripuaren ondoren aseguru-etxekoek kotxea erabat hondaturik zegoela adierazi zuten.
79. Juliari berandu egin zitzaion bere semea liburutik jasotzeko.
80. Ikerketaren emaitzek merkataritza-lokalen ebakuazio-ahaideengan eragina izango dute.

2 List of sentences used in the cue-based recall production task in Experiment 5

1. Edurnek aizkolariari [herriko lehiaketan sari guztiak irabazi dituen] pastela ekarri dio
2. Amaiak armairuaren atal zikin batean ezkutatuta zegoen poltsa andreari eman dio
3. Itziarrek pilotariari diru laguntza gehiegi jaso omen dituen polikiroldegia erakutsi dio
4. Amaiak emakumeari lehen zure lankide bati aipatu diogun formulario luzatu dio
5. Agureak senideari aurreko mendeko euskal arkitekturaren aztarnak dituen baserria utzi dio
6. Sarak kantariari aldizkarietan eta sare sozialetan ikusi duzun argazkia egin
7. Mikelek aitari ordenagailua eta urrutiko kontrolez ibiltzen den kotxea apurtu dio
8. Kazetariak entzuleari pirata ospetsuen abentura ugari kontatzen dituen eleberria gomendatu dio
9. Mekanikoak gidariari hainbat istripuren ondoren erabat hondaturik zegoen furgoneta konpondu dio
10. Zuzendariak aktoreari atzo oso berandu idazten amaitu genuen gidoia irakurri dio
11. Amak irakasleari arratsalde osoan etxeko frigorifikoan izan dugun pastela eraman dio

12. Lehendakariak alkateari Gasteizko eta Bilboko hiritarren babesa daukan plana erakutsi dio
13. Argitaratzaileak itzultzaileari duela bi urte arrakasta izan zuen liburua eskaini dio
14. Ertzainak lapurrari ondoko etxebizitzan odolez zikinduta aurkitu duten pistola kendu dio
15. Editoreak kazetariari aurtengo Euskadi saria irabazteko aukera duen artikulua kritikatu dio
16. Josunek idazkariari neke handiz aurkitu ahal izan duen eserlekua gorde dio
17. Kepak marrazkilariari azken boladan asko aipatua izan den komikia bidali dio
18. Kirmenek gizonari Eibarren eta inguruko herrietan ezaguna den garagadoa atera dio
19. Aitorrek pailazoari nazioartean ospe handia izan omen duen jolasa saldu dio
20. Asierrek neskari aurreko egunean Andoniren dendan ikusi genuen lepokoa erosi dio
21. Iratik taldeari oso jende gutxik gustuko omen duen kanta eskatu dio
22. Zuriñek gizonari gure auzoko emakume guztiek gorroto duten kartela izorratu dio
23. Koldok epaileari entzule guztien artean samina sortu duen iruzkina zuzendu dio
24. Zinegotziak haurrari hiriko kaleetan barrena galduta ibili den txakurra oparitu dio
25. Zientzialariak aurkezleari irrati eta telebista guztietan atera den aurkikuntza azaldu dio
26. Politikariak boluntarioari nazioarteko hedabideetan oso kritikatua izan den erabakia azaldu dio
27. Musikariak abeslariari nire lagunaren iritziz izugarri hunkigarria den balada idatzi dio
28. Zuzendariak ordezkoari gaurko administrazio batzordean goretsia izan den proposamena itzuli dio
29. Tenoreak sopranoari opera munduan lagun asko omen dituen musikaria aurkeztu dio
30. Futbolariak umeari oso garestia eta guztiz bikaina den baloia sinatu dio
31. Begiraleak lankideari familian ia denok aspalditik ezagutzen dugun sekretua aitortu dio
32. Erizainak gaixoari medikuen artean kezka handia piztu duen botika eman dio

3 List of words used in the cue-based recall production task in Experiment 5

1. airea - bilera - zientzia
2. protesta - azoka - jatorria
3. jainkoa - senarra - nobela
4. ordua - txapela - kopurua
5. taberna - zerua - telebista
6. mundua - pistola - bizikleta
7. soldata - zinema - esaldia
8. elurra - gutuna - janaria
9. bakea - usaina - itsasoa
10. larrua - finala - baldintza
11. pareta - liburua - txapelketa
12. sagardoa - gorputza - hasiera
13. arropa - emaitza - familia
14. laguna - hezkuntza - ipuina
15. kantua - eliza - istripua
16. errua - haizea - gizartea
17. kidea - lepoa - osasuna
18. txartela - baloia - sinadura
19. keinua - izeba - maitasuna
20. poema - ogia - aurpegia
21. ibaia - tokia - emaztea
22. osaba - begia - igandea
23. etxea - botila - eguzkia
24. soinua - zerrenda - bazkaria
25. mendia - arima - erregea

26. amona- bidaia - polizia
27. arreba - ezkontza - zenbakia
28. mahaia - bihotza- prezioa
29. hiria - zigorra - eraikuntza
30. lekua - gatazka - irratia
31. ilara - epaia - oparia
32. hizkera - umea - politika

Appendix D

Summary in Spanish - Resumen en castellano

1 Introducción

La preferencia por minimizar la distancia entre los elementos lingüísticos implcados en una dependencia gramatical se ha interpretado como un principio de cognitivo de carácter universal (Ferrer i Cancho, 2004; Futrell et al., 2015b; Hawkins, 1994, 2004a, 2014; Liu, 2008; Temperley, 2008; revisiones del estado de la cuestión, Liu, Xu, & Liang, 2017 y Temperley & Gildea, 2018). Numerosos estudios han defendido que este principio, que favorece las dependencias cortas, modula los órdenes de palabras en el uso lingüístico (Hawkins, 1994, 2004a, 2014), en las gramáticas (Ferrer i Cancho, 2004; Hawkins, 1994; Liu, 2008; Temperley, 2008, entre otros) y en la evolución histórica de las lenguas (Aldai, 2011; Fedzechkina, Chu, & Jaeger, 2018; Tily, 2010).

El objetivo fundamental de esta tesis doctoral ha sido investigar la hipótesis de que la minimización de la longitud de dependencias es un mecanismo activo en la producción del lenguaje, resultado de un principio general de procesamiento (y no de ciertas preferencias estructurales gramaticalizadas propias de una o varias lenguas concretas) y explorar la naturaleza específica de dicho principio, y las condiciones que lo modulan. Para ello, he explorado el efecto de la longitud de las dependencias en las preferencias de producción en lenguas con diferentes propiedades gramaticales (en cuanto a la riqueza de sus sistemas morfológicos de caso y a su órdenes básicos de palabras) y por parte de hablantes nativos y no nativos y con diferente capacidad de memoria de trabajo. He llevado a cabo un total de 5 experimentos de producción, además de 1 tarea de completar oraciones y 2 tareas para medir la memoria de trabajo de los participantes en uno de los experimentos de producción (adaptadas y traducidas de Unsworth, Schrock, & Engle, 2005). En el apartado 1.1 detallo las preguntas en las que he desglosado el tema de investigación de este tesis para poder abordarlo adecuadamente.

Se ha propuesto que la preferencia por las dependencias cortas, atestiguada en un amplio número de gramáticas en lenguas tipológicamente diversas, procede de la convencionalización de dicho principio universal de procesamiento. De ahí la relevancia del objetivo planteado en este tesis, ya que la evidencia procedente de estudios de preferencias lingüísticas en producción es considerablemente más escasa y procede de un número de lenguas mucho más limitado que la que se deriva de estudios de corpus, en los que no se distingue entre preferencias de uso y opciones gramaticales (Ferrer i Cancho, 2004; Futrell et al., 2015b; Gildea and Temperley, 2010; Liu, 2008). En concreto, la evidencia de la preferencia por minimizar dependencias en procesamiento procede de una serie de estudios experimentales en inglés (Stallings, MacDonald, & O'Seaghdh, 1998; Stallings & MacDonald, 2011), japonés (Yamashita & Chang, 2001), coreano (Dennison, 2008) e iraní (Faghiri, Samvelian, & Hemforth, 2014) y una serie pequeños estudios de corpus recabados por Hawkins (1994), que recogen un número considerable de lenguas, pero cuya extensión es pequeña y cuyos resultados no son analizados estadísticamente. Estos trabajos indican que los hablantes prefieren ordenar los elementos cortos antes de los largos en la lenguas de núcleo inicial como el inglés y los largos antes de los cortos en lenguas de núcleo final como el japonés (ver la Figura 1.1 en el Capítulo 1). Ambos órdenes reducen la distancia entre los elementos de una dependencia.

Una de las asunciones básicas en esta tesis doctoral y en gran parte de los trabajos citados en ella es que la minimización de la longitud de dependencias cuando producimos frases responde a consideraciones de eficiencia comunicativa, ya que esta preferencia en producción converge con una tendencia similar en comprensión (Bartek, Lewis, Vasishth, & Smith, 2011; Demberg & Keller, 2008; Gibson, 1998, 2000; Grodner & Gibson, 2005; Levy, Fedorenko, & Gibson, 2013, entre otros). En otras palabras, según ciertos enfoques basados en la teoría de la información, las preferencias lingüísticas son el resultado del equilibrio entre la necesidad de aumentar la probabilidad de que un mensaje determinado se interprete correctamente y la necesidad de reducir el esfuerzo que supone producir dicho mensaje (Bell et al., 2003; Gibson et al., 2013; Horn, 1984; Jaeger, 2010; Kurumada & Jaeger, 2015; Piantadosi, Tily, & Gibson, 2011, Zipf, 1949, entre otros). En concreto, asumo que las dependencias largas suponen un coste cognitivo en comprensión (Gibson, 1998, 2000; Just & Carpenter, 1992; King & Just, 1991; Lewis, Vasishth, & Van Dyke, 2006), pero, a diferencia de otros estudios (Scontras, Badecker, Shank, Lim, & Fedorenko, 2015), no asumo que lo supongan en producción. Los hablantes tratarán de producir órdenes de palabras que reduzcan la longitud de las dependencias de modo que se aseguren de que se da una comunicación robusta al reducir los posibles errores de reconstrucción de la señal por parte del oyente, mientras intentan disminuir en los posible el coste que supone producir los órdenes no canónicos necesarios para reducir la distancia de las dependencias (Myachykov, Scheepers, Garrod, Thompson, & Fedorova, 2011; Myachykov, Thompson, Scheepers, & Garrod, (2013). El equilibrio entre ambas motivaciones conlleva una serie de predicciones compatibles con varios resultados presentados en este trabajo.

1.1 Preguntas de investigación

A continuación listo las principales cuestiones a las que he tratado de contestar en cada uno de los capítulos que conforman esta tesis.

- **Capítulo 2**
  - ¿El orden de los constituyentes en función de su longitud es un efecto de la de accesibilidad conceptual de los sintagmas complejos o el resultado de un principio de procesamiento general, la minimización de la longitud de las dependencias?
  - ¿Emplean las lenguas mecanismos gramaticales específicos como, por ejemplo, la concordancia verbal pluripersonal, para mejorar la eficiencia comunicativa, además de mecanismos generales, comunes a todas ellas, como la minimización de la longitud de dependencias?
- **Capítulo 3**
  - ¿Varía la magnitud de la preferencia por minimizar dependencias entre lenguas en función de la riqueza de la morfología de caso que presenten o de las expectativas derivadas de sus órdenes de palabras básicos?
- **Capítulo 4**
  - Cuando los hablantes bilingües emplean su segunda lengua, ¿transfieren la lineraziación preferida en su lengua materna y/o la frecuencia con la que reducen dependencias en ella?
  - ¿Usan los hablantes no nativos su lengua no dominante de forma diferente que los hablantes nativos, en detrimento de la eficiencia comunicativa y en aras de una mayor facilidad de producción?
- **Capítulo 5**
  - ¿Se debe el principio de minimización de la longitud de dependencias a las limitaciones estructurales del sistema de memoria humano o está relacionado con requerimientos de eficiencia comunicativa?
  - ¿Cómo interactúan las diferencias individuales en la memoria de trabajo con la preferencia por reducir la longitud de las dependencias?

2 Resumen de los capítulos

2.1 CAPÍTULO 2. Mecanismos generales y específicos para la mejora de la eficiencia comunicativa. Evidencia del euskera.

El **Capítulo 2** de esta tesis doctoral contribuye a delinear de forma clara la interacción entre la variación gramatical y la eficiencia comunicativa. En la investigación presentada este capítulo he examinado las preferencias de linearización de los constituyentes de la oración en función de su longitud en euskera mediante un experimento de producción en el que los hablantes ordenaban a su antojo sintagmas de diferente longitud que se les presentaban previamente desordenados. La variable de interés en este estudio era la proporción de órdenes producidos por los participantes en los que la longitud de las dependencias se viera reducida. Este experimento nos permite responder a dos cuestiones concretas. En primer lugar, nos permite contestar a la pregunta de si la tendencia a minimizar la longitud de las dependencias atestiguada en un amplio número de lenguas del mundo está condicionada por requerimientos de eficiencia comunicativa o por la necesidad de facilitar los mecanismos cognitivos involucrados en la producción del lenguaje. En segundo lugar, podemos observar qué tipo de estrategias se emplean para reducir dichas dependencias en lenguas que presentan características tipológicas diferentes a las investigadas hasta ahora como, por ejemplo, la presencia de concordancia verbal rica. En concreto, he investigado si la presencia de un sistema de concordancia verbal pluripersonal como el del euskera, ausente en japonés o coreano, modula las preferencias de producción de ciertos órdenes de palabras cuando algunos de los constituyentes implicados son largos y, por tanto, pueden implicar un coste mayor de procesamiento (Bartek et al. 2011; Demberg & Keller, 2008; Gibson, 1998, 2000; Grodner & Gibson, 2005; Levy et al., 2013, entre otros).

Se ha planteado que la preferencia por ordenar los elementos cortos antes que los largos es una tendencia universal, independiente de propiedades tipológicas, que responde a cuestiones de accesibilidad conceptual: en cualquier lengua, es más fácil acceder a la representación en la memoria de los elementos cortos que de los largos, lo cual hace que requieran menos esfuerzo para su producción (Arnold, Losongco, Wasow, & Ginstrom, 2000, de Smedt, 1994; MacDonald, 2013; Stallings & MacDonald,1998; Stallings & MacDonald, 2011; Wasow, 1997a, 1997b). En cambio, los modelos que defienden el papel de la eficiencia comunicativa en las preferencias de linearización de los constituyentes de diferente longitud predicen que se favorecerá siempre la producción de aquellos órdenes que generen las dependencias más cortas (Futrell, 2017; Hawkins, 1994, 2004a, 2014). Esto implica que en las lenguas de núcleo inicial, como el castellano o el inglés, los elementos cortos precederán a los largos, mientras que, en las lenguas de núcleo final, como el japonés o el euskera, los elementos cortos seguirán a los largos (ver al respecto la Figure 1.1 en el Capítulo 1). Hasta ahora, en lenguas de núcleo final rígidas, que no admiten constituyentes en posición postverbal, se había encontrado una preferencia por los órdenes que minimizaban

dependencias, es decir *largo-antes-de-corto*. Asimismo, en dos experimentos de producción en iraní, en el que solo se investigaron órdenes con el verbo en posición final, se encontró la misma tendencia (Faghiri & Samvelian, 2014; Faghiri et al., 2014). Se ha argüido que dichos resultados pueden explicarse desde la competición entre los efectos de saliencia conceptual y la facilitación de la producción (Chang, 2009; Dennison, 2008; Wasow, 2002, 2013; Yamashita & Chang, 2001). Aunque los sintagamas largos son más difícles de producir, también son conceptualmente más salientes, por lo que en lenguas con libertad en el orden de los constituyentes y con la posibilidad de omitir el sujeto y otros pronombres (lenguas *pro-drop*), los constituyentes más salientes, los más largos, se producirán antes (Wasow, 2002, 2013). También se había sugerido que en lenguas de núcleo final que no permitieran constituyentes postverbales, la preferencia por externalizar los elementos cortos antes de los largos competía con la tendencia a evitar la producción de cláusulas complejas incustradas en el centro de la cláusula (*center-embedding*) (Wasow, 2002), lo que explicaría la falta de órdenes *corto-antes-de-largo*.

El euskera es una lengua de núcleo final, con un orden de palabras más libre que el del japonés o el coreano y que permite desplazar cualquier constituyente a posición postverbal. Las características tipológicas del euskera son, por tanto, claves para testear las dos hipótesis descritas arriba, ya que hacen predicciones opuestas sobre qué órdenes de palabras deberían verse favorecidos en euskera cuando los constituyentes involucrados presentan diferentes longitudes. Los resultados del Experimento 1, presentado en este **Capítulo 2**, indican que en euskera hay una tendencia por ordenar los constituyentes largos antes de los cortos, en contra de lo que se afirma en ciertos trabajos descriptivos sobre esta lengua (Hualde & Ortiz de Urbina, 2003) y de las predicciones de ciertos modelos de producción del lenguaje (Arnold et al., 2000; de Smedt, 1994; MacDonald, 2013; Stallings et al., 1998; Stalllings et al., 2011; Wasow, 1997a, 1997b). La evidencia presentada en este capítulo apoya la idea de que los órdenes especulares de palabras presentes en lenguas que difieren en el parámetro de cabeza (núcleo inicial versus núcleo final) emergen de un principio cognitivo general por minimizar dependencias que favorece la eficiencia comunicativa. Asimismo, los hablantes de euskera tienden a situar el verbo en el medio de la cláusula cuando uno de los constituyentes es largo. Interpreto esta evidencia como prueba de que la morfología verbal, al indicar el rol temático y el caso de los sintagmas que aparecerán más adelante, contribuye a acelerar su procesamiento. Los resultados recogidos en este experimento me llevan a concluir que las lenguas morfolóficamente ricas emplean tanto mecanismos generales que reducen la longitud de las dependencias como específicos que ayudan a incrementear la eficiencia del procesamiento lingüístico.

2.2 CAPÍTULO 3. Diferencias entre lenguas en la preferencia por minimizar dependencias. Evidencia del castellano, el euskera y el polaco.

En el **Capítulo 3**, he investigado si existen diferencias entre el castellano, el euskera y el polaco en cuanto a la magnitud de la preferencia por acortar dependencias, usando exactamente la misma metodología que en el capítulo anterior. La elección del castellano, el euskera y el polaco no es aleatoria. El objetivo principal de la investigación presentada en este capítulo es explorar si la variación encontrada de forma incidental en estudios de corpus de gran tamaño (Futrell et al., 2015b; Gildea & Temperley 2010, Liu, 2008) puede deberse a diferencias en la morfología de caso (pobre en castellano versus rico en polaco y euskera) o a las expectativas sobre la linearización de ciertos constituyentes en función de los órdenes canónicos de dichas lenguas (de núcleo inicial en castellano y polaco versus de núcleo final en euskera). En el marco de los modelos basados en la teoría de la información (Horn, 1984; Shannon, 1948; Zipf, 1949, entre otros), ambas posiblidades implicarían una menor preferencia por la minimización de dependencias. Por un lado, la morfología de caso proporciona pistas claras sobre los elementos que necesitan ser recuperados o integrados en cierto punto concreto de una frase (Tily, 2010; MacWhinney, Bates, & Kliegl, 1984; Futrell, 2017), por lo que recurrir a la minimización de dependencias solo incurriría en redundancia de la señal lingüística. Por otro lado, la probabilidad de que aparezca un elemento lingüístico concreto en un punto de la oración se correlaciona con su coste cognitivo. Cuanto mayor es la expectativa de que aparezca dicho elemento, menor es su coste de procesamiento (Hale, 2001; Levy, 2008). La alta probabilidad de ciertos constituyentes en la cadena hablada implicaría una reducción en su coste, lo que reduciría la necesidad de acortar las dependencias en caso de que estos constituyentes estuvieran implicados en una de ellas.

Los resultados del Experimento 2 en el **Capítulo 3** indican que hay una variación considerablemente amplia entre lenguas tipológicamente diversas y que esta variación no es aleatoria, sino que puede ser explicada en el marco de la teoría de la información. La evidencia presentada en este capítulo sugiere que las diferencias en la magnitud de la preferencia por dependencias cortas está relacionada con la información distribucional generada por el orden básico de palabras propio de cada lengua, y no con diferencias en la riqueza de la morfología de caso. El castellano y el polaco, lenguas ambas de núcleo inicial, siguen un patrón idéntico en las frases transitivas y ditransitivas que testeamos en este experimento y difieren radicalmente del euskera. En las oraciones transitivas, el polaco y el castellano no muestran ningún efecto derivado de la longitud de las dependencias, mientras que el euskera sí lo muestra. En las ditransitivas, el polaco y el castellano presentan una tendencia mucho mayor que el euskera en minimizar dependencias.

Estos resultados no son compatibles con la idea de que la morfología de casos pueda añadir redundancia a la señal lingüística y aumentar la probabilidad de transmitir la información de forma robusta (Tily, 2010; MacWhinney et al., 1984; Futrell, 2017). En este

capítulo defiendo que los cambios en las expectativas sobre los constituyentes que pueden ir apareciendo en la linearización, derivadas de las distribuciones propias de cada lengua, alivian el coste de procesamiento de las dependencias largas (Hale, 2001; Levy, 2008) y reducen, por tanto, la necesidad de acortarlas. Es necesaria más investigación destinada a reducir la posible influencia de otros factores en los resultados obtenidos aquí, entre ellos la posible interacción entre cuestiones de estructure informativa y la longitud de las dependencias en lenguas de núcleo inicial (Aissen, 1999; Arnold et al., 2000; Siewierska, 1993; Wasow, 2002) y el papel de la morofología de caso una vez el efecto de dicha interacción haya podido ser estimado de manera fiable.

Asimismo, en este capítulo replico la tendencia a ordenar los constituyentes cortos antes de los largos en dos lenguas de núcleo inicial tipológicamente lejanas, el castellano y el polaco. La única evidencia experimental previa de esta preferencia procedía de dos estudios sobre el inglés (Stallings et al., 1998; Stallings & MacDonald, 2011). La replicabilidad de los resultados obtenidos en psicologíngüística es clave para garantizar el avance científico del campo, especialmente dada la crisis de replicabilidad que vive la psicología, entre otros campos, y que pone en evidencia la fragilidad de algunos de sus resultados y lo extendido de las llamadas *prácticas de investigación cuestionables* (Open Science Collaboration and others, 2015).

2.3 CAPÍTULO 4. Reducción de la eficiencia comunicativa en el uso no-nativo de la lengua. Evidencia procedente de hablantes bilingües de euskera y castellano

En el **Capítulo 4**, he investigado la preferencia por minimizar la longitud de las dependencias en hablantes bilingües nativos y no nativos de castellano y euskera. Para ello, como en los experimentos reportados en los capítulos anteriores, he analizado los órdenes de palabras producidos por hablantes no-nativos de castellano (Experimento 3) y de euskera (Experimento 4) en función de la longitud de los constituyentes involucrados en los mismos en dos experimentos de producción guiados, y los he comparado con las preferencias mostradas previamente por los hablantes nativos en los Experimentos 1 y 2. Esta comparación resulta infomativa con respecto a la transferibilidad entre lenguas de mecanismos propios de la producción del lenguaje, tales como la minimización de dependencias en hablantes bilingües no nativos. En concreto, el objetivo principal de estos estudios era investigar la afirmación de que incluso los hablantes bilingües altamente competentes en su segunda lengua transfieren a esta la linearización preferida en su lengua nativa (Dennison, 2008). Es decir, los hablantes no nativos de euskera cuya lengua materna sea el castellano tenderán a colocar los constituyentes cortos antes de los largos, tendencia propia de las lenguas de núcleo inicial como el castellano, también cuando hablen en euskera. Igualmente, los hablantes no nativos de castellano cuya lengua materna sea el euskera tenderán a poner los constituyentes largos antes de los cortos, tendencia propia de las lenguas de núcleo final como el euskera, también cuando usen el castellano para comunicarse. Asimismo, quería investigar la

observación, basada en datos de carácter anecdótico, de que los hablantes no nativos de euskera tienden a usar esta lengua de manera menos eficiente que los hablantes nativos, en cuanto a que usan dependencias más largas (Maia, 2015; Zubimendi & Esnal, 1993, entre otros).

El castellano y el euskera, las lenguas usadas en los Experimentos 3 y 4, difieren en dos aspectos de interés, que nos permiten entender de manera más precisa la naturaleza de las consideraciones de eficiencia comunicativa que pueden modular la preferencia por minimizar la longitud de las dependencias en función del orden de palabras propio de la lengua nativa (núcleo inicial versus núcleo final). En primer lugar, estas dos lenguas reducen la longitud de las dependencias a través del uso de linearizaciones opuestas (*largo-antes-de-corto* en euskera y *corto-antes-de-largo* en castellano) (ver los Capítulos 2 y 3). En segundo lugar, el impacto de esta preferencia es mucho mayor en castellano que en euskera (ver el Capítulo 3). Un uso comunicativamente eficiente del lenguaje implica un equilibrio entre la facilitación de la producción del mensaje (mediante el uso mayoritario de órdenes canónicos) y la correcta transmisión del mismo (mediante la minimización de la longitud de las dependencias) (Gibson et al., 2013; Horn, 1984; Zipf, 1949, etc.). Este equilibrio podría verse alterado en el caso de los hablantes no nativos, que favorecerían una producción menos costosa, debido a las mayores exigencias de atención y memoria impuestas por el uso de una lengua no nativa. El uso más o menos eficiente de una segunda lengua por parte de hablantes no nativos es de especial relevancia, no sólo teórica, sino también de índole social, dadas las posibles implicaciones para el empleo de la lengua no dominante, en una comunidad bilingüe como la vasca, en la que conviven dos lenguas de núcleo inicial (castellano o francés) con una de núcleo final (euskera).

Los resultados de los Experimentos 3 y 4 en este capítulo no apoyan la existencia de fenómenos de transferencia del orden superficial preferido en la lengua nativa a la no nativa, al menos en el caso de hablantes bilingües altamente competentes, en contra de lo apoyado en una investigación previa (Dennison, 2008). Los hablantes nativos de castellano no producen órdenes en los que los elementos cortos antecedan a los largos cuando hablan euskera. Igualmente, los hablantes nativos de euskera no producen órdenes en los que los elementos largos precedan a los cortos cuando hablan castellano. Sin embargo, he encontrado evidencia de que los bilingües no nativos altamente competentes y que han sido expuestos a su segunda lengua desde muy temprano, tanto en el caso del castellano como del euskera, difieren de los nativos en cuanto a la frecuencia con la que tienden a la minimización de la longitud de dependencias. Específicamente, los hablantes no nativos reducen la longitud de las dependencias menos frecuentemente que sus homólogos nativos, tanto monolingües como bilingües. Este patrón de resultados sugiere que los hablantes no nativos usan su lengua no dominante de forma menos eficiente que los nativos, ya que emplean más construcciones que incluyen dependencias largas, las cuales resultan cognitivamente más costosas para el oyente. En resumen, la imagen que emerge de los Experimentos 3 y 4 es compatible con la hipótesis

de que las preferencias de los hablantes no nativos pueden deberse a un desequilibrio entre el coste de producción asociado con ciertas construcciones y la transmisión satisfactoria de información. A investigaciones posteriores les corresponderá explorar si se dan cambios cualitativos y cuantitativos en el patrón de desarrollo de la preferencia por reducir la longitud de las dependencias en hablantes bilingües menos competentes o que han comenzado a adquirir su segunda lengua de forma más tardía.

2.4 CAPÍTULO 5. Las consecuencias de las limitaciones de memoria y de las diferencias individuales en la preferencia por minimizar dependencias

La tendencia a minimizar la longitud de las dependencias en las lenguas del mundo se ha relacionado tanto con limitaciones en la capacidad de memoria de trabajo (*Working Memory*) (Futrell, 2017; Scontras et al., Scontras et al., 2015, Liu, 2008; Temperley & Gildea, 2018) como con cosideraciones de eficiencia comunicativa (Futrell et al., 2015b; Gildea & Jaeger, 2015; Hawkins, 1994). En el **Capítulo 5**, he abordado por primera vez hasta donde yo sé, las causas subyacentes a la minimización de dependencias en producción. Para ello, he llevado a cabo un experimento de doble cometido (*dual-task*), en el que los participantes tenían que mantener una carga menor (1 palabra) o mayor (3 palabras) en memoria al tiempo que producían frases que incluían constituyentes de diferente longitiud (corto versus largo). Asimismo, he incluido dos tareas, una de atención de lectura (*Reading span task*) y otra de atención aritmética (*Operation span task*) (Unsworth et al., 2005). Estas tareas sirven para medir las diferencias individuales con respecto a la capacidad de memoria a corto plazo de los participantes, hablantes nativos de euskera. He medido la proporción de órdenes en los que los participantes minimizaban la longitud de las posibles dependencias en función de la carga de memoria que se les presentaba y de su propia capacidad de memoria a corto plazo. Mi objetivo en esta investigación ha sido obtener información relevante sobre la dirección de la relación entre las demandas impuestas a la memoria de trabajo y la preferencia por dependencias cortas, para lo que los investigadores han ofrecido dos posibilidades diferentes, como he mencionado más arriba. Por una parte, se ha sugerido que la preferencia por las dependencias cortas emege como resultado de la interacción entre la incrementalidad propia del sistema de producción del lenguaje y la capacidad limitada del sistema de memoria humano (Futrell, 2017; Jaeger, 2010; Liu, 2008). Por otra, la convergencia entre la preferencia por dependencias cortas en comprensión y en producción ha llevado a otros investigadores a plantear la posibilidad de que dicha preferencia en la producción del lenguaje sea el resultado de consideraciones de eficiencia comunicativa (Fedzechkina, Jaeger, & Newport, 2012; Fedzechkina et al. 2018; Ferrer i Cancho, 2004; Hawkins, 1994, 2004a, 2014).

A continuación describo los resultados obtenidos en el Experimento 5, reportado en dicho capítulo. En primer lugar, la preferencia por la minimización de dependencias aumenta a medida que crece la capacidad de memoria de trabajo de los participantes, medida por la

tarea de atención a la lectura (*Reading span task*). Este resultado no pueden explicarse de manera directa por parte de ninguna teoría que defienda que la preferencia por minimizar la longitud de las dependencias se debe al coste cognitivo impuesto en memoria por parte de las dependencias largas frente a las cortas, ya que son los participantes con menor memoria de trabajo los que tienden a producir dependencias más largas. Este resultado es, además, compatible con la asunción, clave en esta tesis doctoral, de que la preferencia por minimizar dependencias no se debe al coste de producción, sino de comprensión de las mismas (Gibson, 1998, 2000; Just and Carpenter, 1992; King and Just, 1991; Lewis et al., 2006). Sin embargo, no encuentro ninguna evidencia de correlación entre la memoria de trabajo, medida por la tarea de atención aritmética (*Operation span task*), y la producción de órdenes que generen dependencias más cortas, lo que sugiere o bien una falta de poder estadístico en mi diseño experimental o bien la necesidad de repensar el papel otorgado a la habilidad lingüística o a los procesos de comprensión en cualquier modelo que trate de dar cuenta de la preferencia por minimizar dependencias en producción. En segundo lugar, en la tarea de doble cometido (*dual-task*), los participantes de este experimento empleaban órdenes con dependencias cortas de manera más frecuente cuando tenían una carga concurrente de memoria pequeña, aunque únicamente cuando habían visto dichos órdenes previamente en la presentación escrita de los estímulos. Este resultado sugiere nuevamente que existe algún tipo de relación entre los procesos de memoria y la preferencia por minimizar dependencias. Sin embargo, como en el caso anterior, este patrón de resultados solo podría explicarse dentro de un marco que incluya una interrelación más compleja entre el procesamiento lingüístico y las caracteristicas estructurales del sistema de memoria humano que la planteada hasta ahora.

3 Conclusiones generales

Uno de los aspectos metodólogicos claves en esta tesis es la adopción de un punto de vista croslingüístico. El objetivo científico de la psicolingüística es descubrir los principios y mecanismos cognitivos universales en el procesamiento del lenguaje, comunes a todas las lenguas y hablantes. Incluir lenguas tipológicamente distantes nos permite no solo una mayor confianza en los resultados obtenidos, actualmente sesgados por la omnipresencia del inglés como objetivo privilegiado de estudio (Jaeger & Norcliffe, 2009; Norcliffe, Harris, and Jaeger,2010), sino también percatarnos más fácilmente de huecos y problemas presentes en las principales teorías de nuestro campo, que pudieran haber pasado desapercibidos de otra manera. Esta tesis pretende contribuir en ambas direcciones al estado actual de nuestros conocimientos sobre la preferencia por dependencias cortas en la producción del lenguaje, tal y como se desprende del listado que presento a continuación, en el cual resumo lo que considero las principales contribuciones de esta tesis, descritas en la sección anterior de manera más prolija.

1. Lenguas tan diversas como el castellano, el euskera o el polaco presentan una clara tendencia a reducir la longitud de las dependencias (ver los Capítulos 2 y 3). Esta evidencia se suma a la aportada por lenguas como el inglés, el coreano o el japonés, lo que apunta de forma sólida a que la minimización de las dependencias es un principio ubicuo en el procesamiento del lenguaje humano. En resumen, lenguas poco estudiadas en el ámbito de la psicolingüística como el castellano, el euskera y el polaco, indican que, cuando existe más de una alternativa posible para transmitir el mismo contenido, hay una clara preferencia por parte de los hablantes de producir aquellos órdenes de palabras que reducen la distancia entre elementos ligados en una dependencia.

   (a) Los datos aportados por el euskera son de particular importancia (ver el Capítulo 2), ya que nos permiten concluir que la variación en las direcciones de linearización de constituyentes de distinta longitud en lenguas de núcleo inicial y núcleo final, originalmente interpretadas como diferencias específicas en relación con fenónemos de accesibilidad conceptual en lenguas con diferentes especificaciones gramaticales, es, en realidad, indicativa un sesgo cognitivo general por reducir la distancia de las dependencias sintácticas.

   (b) Presento evidencia experimental en castellano y polaco, que se suma a la ya existente en inglés en apoyar la hipótesis de que las lenguas de núcleo inicial los consituyentes cortos tienden a preceder a los largos(ver Capítulo 3). Estos resultados son relevantes en la teorización sobre los posibles mecanismos generales de producción del lenguaje, ya que hasta ahora solo se contaba con datos experimentales procedentes de una única lengua, el inglés

2. Ofrezco evidencia de que la tendencia a acortar la longitud de las dependencias varía de forma consistente entre lenguas (ver Capítulo 3) y entre hablantes (ver Capítulo 4). He tratado de reconciliar este resultado con asunciones propias de teoría de la información.

   (a) He sugerido que este nivel óptimo de eficiencia comunicativa podría obtenerse a través de estrategias específicas de ciertas lenguas, como la concordancia pluripersonal en euskera (ver el Capítulo 2 y la propuesta de Ueno (2009), Hiranuma, 1999 o Pastor and Laka, 2013), o de índole general como la construccion de expectativas basadas en el input proporcionado por cada lengua (ver el Capítulo 3).

   (b) La longitud de las dependencias varía de forma consistente también entre hablantes en función de su competencia o exposición lingüística a una segunda lengua y de su capacidad de memoria (ver Capítulos 4 y 5).

3. Abordo las posibles causas que subyacen a la preferencia por la minimización de la longitud dependencias en las lenguas del mundo. Específicamente, me he centrado en explorar el papel que juegan las características del sistema de memoria humano en modular esta tendencia (ver el Capítulo 5). No he encontrado evidencia clara que apoye la idea de que la preferencia por dependencias cortas se debe a limitaciones en la

memoria de trabajo. Mis resultados sugieren más bien que las consideraciones de eficiencia comunicativa constituyen una línea de investigación más prometedora.

4 Futuras líneas de investigación

Se ha afirmado que la minimización de la longitud de las dependencias influencia las gramáticas de las lenguas del mundo y cómo estas evolucionan a lo largo de la historia (Aldai, 2011; Fedzechkina et al. 2018; Futrell et al., 2015b; Hawkins, 1994, 2004a, 2014; Liu, 2008; Tily, 2010). Para que esto pueda suceder, es primero indispensable que la preferencia por reducir la longitud de las dependencias se de en las preferencias de uso de la lengua, que posteriormente pueden verse gramaticalizadas (Hawkins, 1994). Es, por tanto, esencial que se tenga una idea más precisa de la naturaleza de este preferencia de procesamiento y de su interacción con otros factores. El trabajo presentado en esta tesis supone un primer paso en esta dirección, pero queda mucho por hacer. En este resumen mencionaré dos líneas de investigación que considero fundamentales.

En primer lugar, si en algunas lenguas, como en el euskera, la preferencia por los órdenes más eficientes con dependencias cortas es mínima, no está claro cómo es posible que la minimización de la longitud de dependencias pueda convencionalizarse y afectar a sus gramáticas de una forma tan profunda como se ha afirmado en la literatura (Ferrer i Cancho, 2006; Futrell et al., 2015b; Liu, 2008; Temperley, 2007, entre otros). Es necesario que se obtenga una estimación más fiable de la preferencia por minimizar dependencias durante el procesamiento lingüístico. Para ello, es crucial que se esclarezca la relación entre la preferencia por acortar dependencias y la estructura de la información y otras cuestiones de índole pragmática, cuyos los efectos en lenguas de núcleo inicial como el castellano o el inglés están correlacionados (Wasow, 2002). Asismimo, es necesario que se examine en profundidad el papel del grado de libertad en el orden de palabras propio de cada lengua y su interrelación con la preferencia por dependencias cortas. Además, es fundamental que se tenga en cuenta que la magnitud de la tendencia a minimizar dependencias varía de forma consistente no solo entre lenguas, sino también entre hablantes, lo que supone un reto para cualquier enfoque que quiera establecer una conexión causal entre preferencias de procesamiento y propiedades gramáticales. Por último, es importante investigar si la preferencia por minimizar dependencias también guía la adquisición del lenguaje, y no únicamente el uso adulto de la lengua.

En segundo lugar, es de crucial importancia que se investigue hasta qué punto se correlacionan las preferencias por dependencias cortas en comprensión y en producción en las lenguas testeadas hasta ahora. En inglés hay una gran cantidad de estudios que muestran que las dependencias cortas facilitan el procesamiento (Bartek et al. 2011; Demberg & Keller, 2008;

Gibson, 1998, 2000; Grodner & Gibson, 2005; Levy et al., 2013, entre otros). Sin embargo, efectos llamados de anti-localidad (*anti-locality*), donde las dependencias largas facilitan (o al menos no dificultan) el procesamiento, se han reportado en lenguas como el hindi, el alemán o el japonés (Konieczny & Döring, 2003; Nakatani2010; Vasishth & Drenhaus, 2011). A este respecto, no hay datos disponibles sobre el castellano, el euskera o el polaco. Futuros trabajos tendrán que abordar cómo de tolerables son las dependencias largas en estas tres lenguas (y en otras) y en qué grado se correlacionan con las preferencias en producción. En cualquier caso, es sugestivo que sean las lenguas de núcleo final aquellas en las que hasta ahora se ha encontrado una tendencia a presentar efectos de anti-localidad en comprensión, coincidiendo con el efecto más débil de la preferencia por minimizar dependencias en producción que hemos encontrado en euskera (ver Capítulo 3) y que parece aparecer también en otras lenguas de núcleo final (Futrell et al., w2015b).

Bibliography

Abney, S. P. (1987). *The English Noun Phrase in its sentential aspect* (Unpublished doctoral dissertation, Cambridge, MA: Massachusetts Institute of Technology).

Aduriz, I., Aranzabe, M. J., Arriola, J. M., Atutxa, A., Dıaz de Ilarraza, A., Garmendia, A., & Oronoz, M. (2003). Construction of a Basque dependency treebank. *Proceedings of the 2nd Workshop on Treebanks and Linguistics Theories (TLT)*, 201–204.

Aissen, J. (1999). Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory*, *17*, 673–711.

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory*, *21*(3), 435–483.

Aldai, G. (2011). Wh-questions and SOV languages in Hawkins' (2004) theory: Evidence from Basque. *Linguistics*, *49*(5), 1079–1135.

Altube, S. (1929). *Erderismos*. Bilbo: Cinsa.

Amuriza, X. (2012). *Zazpi ebidentzia birjaiotzarako*. Bilbo: Lanku.

Anderson, J. R. (1990). *The adaptive character of thought*. Hilsdale, NJ: Erlbaum.

Anttila, A., Adams, M., & Speriosu, M. (2010). The role of prosody in the English dative alternation. *Language and Cognitive Processes*, *25*(7-9), 946–981.

Aristegieta, X. (2009). Euskarazko testuen komunikagarritasun-problema larriak. *Senez*, *37*(3), 103–141.

Arnold, J. E., Losongco, A., Wasow, T., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, *76*, 28–55.

Arnold, J. E., Wasow, T., Asudeh, A., & Alrenga, P. (2004). Avoiding attachment ambiguities: The role of constituent ordering. *Journal of Memory and Language*, *51*(1), 55–70.

Audacity Team. (2016). Audacity(R): Free Audio Editor and Recorder [Computer software].

Aylett, M. & Turk, A. (2004). The Smooth Signal Redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic Prominence, and duration in spontaneous speech. *Language and Speech*, *47*(1), 31–56.

Baayen, R. H. (2004). Statistics in psycholinguistics: A critique of some current gold standards. *Mental lexicon working papers*, *1*(1), 1–47.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Oxford University Press.

Baddeley, A. D. & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The Psychology of learning and motivation* (Vol. 8, pp. 47–89). New York, NY: Academic Press.

Bartek, B., Lewis, R. L., Vasishth, S., & Smith, M. R. (2011). In search of on-line locality effects in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1178–1198.

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1. 1-7.

Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, *11*(3), 245–299.

Behaghel, O. (1909/10). Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen*, *25*, 110–42.

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *The Journal of the Acoustical Society of America*, *113*(2), 1001–1024.

Bickel, B. (2015). Distributional typology: Statistical inquiries into the dynamics of linguistic diversity. In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Linguistics Analysis (2nd edition)* (pp. 901–923). Oxford, UK: Oxford University Press.

Bickford, J. A. (1985). Spanish clitic doubling and levels of grammatical relations. *Lingua*, *65*(3), 189–211.

Blake, B. J. (1994). *Case*. Cambridge, UK: Cambridge University Press.

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, *89*, 1–47.

Bock, J. K. & Irwin, D. E. (1980). Regulating mental energy: Performance units in language production. *Journal of Verbal Learning and Verbal Behavior*, *19*, 467–484.

Bock, J. K. & Levelt, J. W. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945–984). Orlando, FL: Academic Press.

Bock, J. K. & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*, 99–127.

Bock, J. K. & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, *21*(1), 47–67.

Bolinger, D. & Gertsman, L. (1957). Disjuncture as a clue to constructs. *Word*, *13*, 246–255.

Boston, M. F., Hale, J., Kliegl, R., Patil, U., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research*, *2*(1), 1–12.

Branigan, H. P., Pickering, M. J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, *118*(2), 172–189.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, H. (2007). Predicting the dative alternation. In G. Boume, I. Kraemer, & J. Zwarts (Eds.), *Cognitive Foundations of Interpretation* (pp. 69–94). Amsterdam, NL: Royal Netherlands Academy of Science.

Bresnan, J. & Ford, M. (2010). Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language*, *86*(1), 168–213.

Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. John Benjamins.

Calvo, N., Ibañez, A., & García, A. M. (2016). The impact of bilingualism on working memory: A null effect on the whole may not be so on the parts. *Frontiers in Psychology*, *7*, 265.

Caplan, D. & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*(1), 265.

Chang, F. (2009). Learning to order words: A connectionist model of heavy NP shift and accessibility effects in Japanese and English. *Journal of Memory and Language*, *61*(3), 374–397.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272.

Choi, H.-W. (2007). Length and order: A corpus study of Korean dative-accusative construction. *Discourse and Cognition*, *14*(3), 207–227.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Christiansen, M. H. (1974). Using artificial language learning to study language evolution: Exploring the emergence of word order universals. In L. Ghadakpour & J. Desalles (Eds.), *The evolution of language: Third International Conference* (pp. 45–48). Paris: École National Superieure des Telecommunications.

Christiansen, M. H. & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*(5), 489–509.

Christiansen, M. H. & Chater, N. (2016). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*.

Clark, E. V. & Clark, H. H. (1979). When nouns surface as verbs. *Language*, *55*, 287–299.

Clark, H. H. (1973). The Language-as-fixed-effect fallacy: A critique of language statistics in psyhological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359.

Clark, H. H. & Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in psychology*, *9*, 287–299.

Collins, M. (2003). Head-driven statistical models for natural language parsing. *Computations Linguistics*, *29*, 589–638.

Comrie, B. (1989). *Language Universals and Linguistic Typology (2nd edition)*. Chicago, ILL: University of Chicago Press.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review*, *12*(5), 769–786.

Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and Brain Sciences*, *24*(01), 154–176.

Croft, W. A. (1990). *Typology and Universals*. Cambridge, UK: Cambridge University Press.

Croft, W. A. (2001). Functional approaches to grammar. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social and Behavioral Sciences* (Vol. 12, pp. 6323–6330). Oxford, UK: Elsevier Sciences.

Culbertson, J. & Adger, D. (2014). Language learners privilege structured meaning over surface frequency. *Proceedings of the National Academy of Science*, *111*, 5842–5847.

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, *122*(3), 306–329.

Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research*, *20*, 4445–464.

Daneman, M. & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

Daneman, M. & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, *25*(1-18), 450–466.

Daneman, M. & Merikle, P. M. (1986). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review*, *3*(4), 422–433.

De Rijk, R. P. G. (1969). Is Basque an SOV language? *Fontes Linguae Vasconum*, 319–351.

De Rijk, R. P. G. (2007). *Standard Basque: A progressive grammar*. MIT Press.

de Marneffe, M.-C., Grimm, S., Arnon, I., Kirby, S., & Bresnan, J. (2012). A statistical model of the grammatical choices in child production of dative sentences. *Language and Cognitive Processes*, *27*(1), 25–61.

de Smedt, K. (1994). Parallelism in incremental sentence generation. In G. Adriaes & U. Hahn (Eds.), *Parallelism in Natural Language Processing (* (pp. 421–447). Norwood, NJ: Able.

Dell, G. S. & O'Seaghdha, P. G. (1994). Inhibition in interactive activation models of linguistic selection and sequencing. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory Processes in Attention, Memory and Language* (pp. 409–453). San Diego, CA: Academic Press.

Demberg, V. & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, *109*(2), 193–210.

Demonte, V. (1995). Dative alternation in Spanish. *Probus*, (7), 5–30.

Dennison, H. (2008). Universal versus language-specific conceptual effects on shifted word-order production in Korean: Evidence from bilinguals. *Working Papers in Linguistics: University of Hawaii at Manoa*, *39*(2), 1–16.

Department of Education, L. P. & Culture. (2013a). *Fifth Sociolinguistic Survey 2011*. Vitoria-Gasteiz: Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia / Servicio Central de Publicaciones del Gobierno Vasco.

Department of Education, L. P. & Culture. (2013b). *Talking Pupils. The Arrue Porject 2011•*. Vitoria-Gasteiz: Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia / Servicio Central de Publicaciones del Gobierno Vasco.

Dryer, M. S. (1980). The positional tendencies of sentential Noun Phrases in Universal Grammar. *Canadian Journal of Linguistics*, *25*(1), 81–138.

Dryer, M. S. (1992). The Greenbergian word order correlations. *Language*, *68*(1), 81–138.

Dryer, M. S. (2002). Case distinctions, rich verb agreement, and word order type (comments on Hawkins' paper). *Theoretical Linguistics*, *28*(2), 151–158.

Dryer, M. S. (2011). The evidence for word order correlations: A response to Dunn, Greenhill, Levinson's and Gray's paper in Nature. *Linguistic Typology*, *15*(2), 335–380.

Dunn, M., Greenhill, S. J., Levinson, S. C., & Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, *473*(7345), 79–82.

Dye, M., Milin, P., Futrell, R., & Ramscar, M. (2018). Alternative solutions to a language design problem: The role of adjectives and gender marking in efficient communication. *Topics in Cognitive Science*, *10*(1), 209–224.

Eisner, J. & Smith, N. (2005). Parsing with soft and hard constraints on dependency length. In *Proceedings of the 9th International Workshop on Parsing Technology (IWPT 2005* (pp. 30–41).

Elordieta, A. (2001). *Verb movement and constituent permutation in Basque*. LOT.

Erdocia, K., Laka, I., A., Mestres-Miss'e, & Rodríguez-Fornells, A. (2009). Syntactic complexity and ambiguity resolution in a free word order language: Behavioral and electrophysiological evidences from Basque. *Brain and Language*, *109*, 1–17.

Evans, N. & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*(5), 429–48.

Faghiri, P. & Samvelian, P. (2014). Constituent ordering in Persian and the weight factor. *Empirical Issues in Syntax and Semantics 10*, 215–232.

Faghiri, P., Samvelian, P., & Hemforth, B. (2014). Accessibility and word order: The case of ditransitive constructions in Persian. In *21th International Conference on Head-Driven Phrase Structure Grammar* (pp. 217–237).

Farmer, T. A., Fine, A. B., Misyak, J. B., & Christiansen, M. H. (2017). Reading span task performance, linguistic experience, and the processing of unexpected syntactic events. *Quaterly Journal of Experimental Psychology*, *70*(3), 413–433.

Farmer, T. A., Misyak, J. B., & Christiansen, M. H. (2012). Individual differences in sentence processing. In M. F. J. M. J. Spivey & K. McRae (Eds.), *Cambridge Handbook of Psycholinguistics* (pp. 353–364). Cambridge, UK: Cambridge University Press.

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language*, *54*(4), 541–553.

Fedorenko, E., Gibson, E., & Rohde, D. (2007). The nature of working memory in linguistic, arithmetic and spatial integration processes. *Journal of Memory and Language*, *56*(2), 246–269.

Fedzechkina, M., Chu, B., & Jaeger, T. F. (2018). Human processing shapes language change. *Psychological Science*, *29*(1), 72–82.

Fedzechkina, M., Jaeger, T. F., & Newport, E. L. (2012). Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences*, *109*(44), 17897–17902.

Fedzechkina, M., Jaeger, T. F., & Newport, E. L. (2013). Communicative biases shape structures of newly acquired languages. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 430–435). Cognitive Science Society.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, *30*(2), 210–233.

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*(1), 11–15.

Ferreira, F. & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*(1), 57–84.

Ferreira, V. S. (1996). Is it better to *give* than to *donate*? Syntactic flexibility in language production. *Journal of Memory and Language*, *35*(5), 724–755.

Ferreira, V. S. & Yoshita, H. (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research*, *32*, 669–692.

Ferreira, V. S. & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, *40*(4), 296–340.

Ferrer i Cancho, R. (2004). Euclidean distance between syntactically linked words. *Physical Review E*, *70*(5 Pt 2), 056135.

Ferrer i Cancho, R. (2005). Zipf's law from a communicative phase transition. *The European Physical Journal B-Condensed Matter and Complex Systems*, *47*(3), 449–457.

Ferrer i Cancho, R. (2006). Why do syntactic links not cross? *EPL (Europhysics Letters)*, *76*(6), 1228.

Ferrer i Cancho, R. & Díaz-Guilera, A. (2007). Zipf's law from a communicative phase transition. *Journal of Statistical Mechanics: Theory and Experiment*, (P06009), 449–457.

Ferrer i Cancho, R. & Liu, H. (2014). The risks of mixing dependency lengths from sequences of different length. *Glottotheory*, *5*(2), 143–155.

Ferrer i Cancho, R. & Solé, R. (2003). Least effort and the origins of scaling in human language. *Proceedings of the National Academy of Sciences*, *100*(3), 788.

Ferrer i Cancho, R. (2015). The placement of the head that minimizes online memory: A complex systems approach. *Language Dynamics and Change*, *5*(1), 114–137.

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*, *8*(10), 143–155.

Flores, M. (2004). OD preposicional y OI duplicado: Sobre la vinculación de dos cambios. In J. J. Bustos & J. L. Girón (Eds.), *Actas del VI Congreso Internacional de Historia de la Lengua Española: Madrid, 29 de septiembre-3 octubre 2003* (pp. 667–676). Arcos Libros, UCM and AHLE.

Fodor, J. (1983). *The modularity of the mind*. Cambridge, MA: MIT Press.

Forster, K. I. & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers*, *35*(1), 116–124.

Francis, J. E. (2006). Grammatical weight and relative clause extraposition in English. *Cognitive Linguistics*, *101*(21), 35–74.

Fraundorf, S. H. & Jaeger, T. F. (2016). Readers generalize adaptation to newly-encountered dialectal structures to other unfamiliar structures. *Journal of Memory and Language*, *91*, 28–58.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies* (Journal Article).

Frazier, L. (1985). Syntactic complexity. In L. K. David R. Dowty & A. M. Zwicky (Eds.), *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives* (pp. 129–189). Cambridge University Press.

Frazier, L. (1990). Exploring the architecture of the language-processing system. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 383–408). MIT Press.

Frazier, L. & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*(4), 291–325.

Frenck-Mestre, C. (2005). Ambiguities and anomalies: What can eye movements and event related potentials reveal about second language sentence processing? In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of Bilingualism. Psycholinguistic approaches* (pp. 268–284). Oxford University Press.

Futrell, R. (2017). *Memory and locality in natural language* (Doctoral dissertation, PhD thesis. Cambridge, MA: Massachusetts Institute of Technology).

Futrell, R. & Levy, R. (2017). Noisy-context surprisal as a human sentence processing cost model. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics* (Vol. 1, pp. 688–698).

Futrell, R., Mahowald, K., & Gibson, E. (2015a). *ArXiv e-prints*. arXiv: 1510.00436 [cs.CL]

Futrell, R., Mahowald, K., & Gibson, E. (2015b). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences*, *112*(33), 10336–10341.

Futrell, R., Mahowald, K., & Gibson, E. (2015c). Quantifying word order freedom in dependency corpora. In *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)* (pp. 91–100).

Gahl, S. & Garnsey, S. M. (2004). Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language*, *80*, 748–775.

Gennari, S. P., Mirkovic, J., & MacDonald, M. C. (2012). Animacy and competition in relative clause production: A cross-linguistic investigation. *Cognitive Psychology*, *65*(2), 141–76.

Genzel, D. & Charniak, E. (2003). Variation of entropy and parse trees of sentences as a function of the sentence number. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing, ACL* (pp. 65–72).

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1–76.

Gibson, E. (2000). The Dependency Locality Theory: A distance-based theory of linguistic complexity. *Image, Language, Brain*, 95–126.

Gibson, E. & Fedorenko, E. (2010). Weak quantitative standards in Linguistics research. *Trends in Cognitive Sciences*, *14*(6), 233–234.

Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of crosslinguistic word-order variation. *Psychological Science*, *24*(7), 1079–88.

Gibson, E., Piantadosi, S. T., & Fedorenko, E. (2010). Quantitative methods in syntax/semantics research: A response to Sprouse and Almeida (2013). *Language and Cognitive Processes*, *28*(3), 229–240.

Gildea, D. & Jaeger, T. F. (2015). Human languages order information efficiently. *ArXiv e-prints*. arXiv: 1510.02823 [cs.CL]

Gildea, D. & Temperley, D. (2010). Do grammars minimize dependency length? *Cognitive Science*, *34*(2), 286–310.

Givón, T. (1991). Isomorphism in the grammatical code: Cognitive and biological considerations. *Studies in Language*, *1*(15), 85–114.

Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, *105*(27), 9163–9168.

Gómez-Rodríguez, C. (2016). Commentary on Natural Language Processing and the Now-or-Never Bottleneck. *Behavioral and Brain Sciences*, *39*, 30.

Gómez-Rodríguez, C. (2017). On the relation between dependency distance, crossing dependencies, and parsing. Comment on "Dependency distance: a new perspective on syntactic patterns in natural languages" by Haitao Liu et al. *Physics of Life Reviews*, *21*, 200–203.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, *51*(1), 97–114.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological science*, *13*(5), 425–430.

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. *Universals of language*, *2*, 73–113.

Gregory, M. & Michaelis, L. (2001). Topicalization and left-dislocation: a functional opposition revisited. *Journal of Pragmatics*, *33*(11), 1665–1706.

Grodner, D. & Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentenial complexity. *Cognitive Science*, *29*(2), 261–290.

Gulordava, K. & Merlo, P. (2015). Diachronic trends in word order freedom and dependency length in dependencyannotated corpora of Latin and ancient Greek. In *Proceedings of the third international conference on dependency linguistics (Depling 2015)* (pp. 121–130).

Gulordava, K. & Merlo, P. (2016). Multi-lingual dependency parsing evaluation: A large-scale analysis of word order properties using artificial data. *Transactions of the Association for Computational Linguistics*, *4*, 343–356.

Haider, H. (in press). VO/OV-base ordering in Germanic languages. In R. Page & M. T. Putnam (Eds.), *Cambridge Handbook of Germanic linguistics* (pp. 3–21). Cambridge University Press.

Hakuta, K. (1981). Grammatical description versus configurational arrangement in language acquisition: The case of relative clauses in Japanese. *Cognition*, *9*, 197–236.

Hakuta, K. & d'Andrea, D. (1992). Some properties of bilingual maintenance and loss in Mexican background high-school students. *Applied Linguistics*, *13*(1), 72–99.

Hale, J. (2001). A Probabilistic Early Parser As a Psycholinguistic Model. In *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics on Language Technologies* (pp. 1–8). NAACL '01. Pittsburgh, Pennsylvania: Association for Computational Linguistics.

Hartsuiker, R. J. & Barkhusen, P. N. (2006). Language production and working memory: The case of subject-verb agreement. *Language and Cognitive Processes*, *21*(1-3), 181–204.

Haspelmath, M. (1999). Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft*, 180–205.

Haspelmath, M. (2008). Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics*, *19*(1), 1–33.

Haviland, S. E. & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, *13*(5), 512–521.

Hawkins, J. A. (1983). *Word order universals*. New York, NY: Academic Press.

Hawkins, J. A. (1994). *A performance theory of order and constituency*. New York, NY: Cambridge University Press.

Hawkins, J. A. (2004a). *Efficiency and complexity in grammars*. New York, NY: Oxford University Press.

Hawkins, J. A. (2004b). Language complexity as an evolving variable. In G. D. Sampson G & T. P (Eds.), *Constituent Order in the Languages of Europe* (pp. 158–175). New York, NY: Oxford University Press.

Hawkins, J. A. (2014). *Cross-linguistic variation and efficiency.* Oxford University Press.

Heringer, H. J., Strecker, B., & Wimmer, R. (1980). *Syntax*. W. Fink.

Hiranuma, S. (1999). Syntactic difficulty in English and Japanese: A textual study. *UCL Work. Papers Linguist*, *11*, 309–322.

Hockett, C. F. (1960). The Origin of Speech. *Scientific American*, *203*(3), 88–97.

Hoffmann, C. (1999). Word order and the principle of "Early Immediate Constituents"(EIC). *Journal of Quantitative Linguistics*, *6*(2), 108–116.

Horn, L. R. (1984). Towards a new taxonomy for pragmatic inference: Q-based and R-based implicature. *Meaning, form, and use in context: Linguistic applications*, *6*(11), 11–42.

Hualde, J. I. & Ortiz de Urbina, J. (2003). *A Grammar of Basque*. Berlin: Mouton de Gruyter.

Hudson Kam, C. L. & Newport, E. L. (2016). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, *59*(1), 30–66.

Hudson, R. (1991). *English Word Grammar*. Oxford, UK: Wiley-Blackwell.

Hudson, R. (1995). Measuring syntactic difficulty. *Manuscript, University College, London*.

Hudson, R. (2017). Cross-language diversity, head-direction and grammars Comment on "Dependency distance: A new perspective on syntactic patterns in natural languages" by Haitao Liu et al. *Physics of Life Reviews*, *21*, 204–206.

Humphreys, G. F., Mirković, J., & Gennari, S. P. (2016). Similarity-based competition in relative clause production and comprehension. *Journal of Memory and Language*, *89*, 200–221.

Husain, S., Vasishth, S., & Srinivasan, N. (2014). Strong expectations cancel locality effects: Evidence from Hindi. *PloS One*, *9*(e100986).

Hwang, H. & Kaiser, E. (2014). Having a syntactic choice is not always better: the effects of syntactic flexibility on Korean production. *Language, Cognition and Neuroscience*, *29*(9), 1115–1131.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, *61*(1), 23–62.

Jaeger, T. F. (2011). Corpus-based research on language production information density and reducible subject relatives. In E. M. Bender & J. E. Arnold (Eds.), *Language from a cognitive perspective: Grammar, usage, and processing. Studies in honor of Tom Wasow* (pp. 161–197). Standford, CA: CSLI Publications.

Jaeger, T. F. (2013). Production preferences cannot be understood without reference to communication. *Frontiers in Psychology*, *4*, 230.

Jaeger, T. F. & Buz, E. (2017). Signal reduction and linguistic encoding. In E. M. Fernández & H. S. Cairns (Eds.), *Handbook of Psycholinguistics* (pp. 38–81). Wiley-Blackwell.

Jaeger, T. F. & Norcliffe, E. (2009). The cross-linguistic study of sentence production. *Language and Linguistics Compass*, *3*(4), 866–887.

Jaeger, T. F. & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*(1), 57–83.

Jaeger, T. F. & Tily, H. (2010). On language 'utility': processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(3), 323–35.

Jaeggli, O. (1982). *Topics in Romance Syntax (Studies in generative grammar)*. Foris Publications.

Jäger, L., Chen, Z., Li, Q., Lin, C.-J. C., & Vasishth, S. (2015). The subject-relative advantage in Chinese: Evidence for expectation-based processing. *Journal of Memory and Language*, *79-80*(3), 97–120.

Jiang, J. & Liu, H. (2015). The effects of sentence length on dependency distance, dependency direction and the implications–Based on a parallel English–Chinese dependency treebank. *Language Sciences*, *50*, 93–104.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, *59*, 193–224.

Just, M. A. & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149.

Kaltzakorta, M. (2007). *Prosa komunikagarriago egiten (zenbait proposamen)*. Bilbao: UEU.

Keenan, E. L. (1979). On Surface Form and Logical Form. *Studies in the Linguistic Sciences*, *8*(2), 163–203.

Kellog, R. T. (2004). Working memory components in written sentence generation. *American Journal of Psychology*, *117*(1), 341–361.

Kempe, V. & MacWhinney, B. (1999). Processing of morphological and semantic cues in Russian and German. *Language and Cognitive Processes*, *14*, 129–171.

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences*, *22*(2), 154–169.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, *2*(1), 15–47.

King, J. & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of memory and language*, *30*(5), 580–602.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of*

the National Academy of Sciences, *105*(31), 10681–10686. eprint: http://www.pnas.org/content/105/31/10681.full.pdf

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.

Kizach, J. (2012). Evidence for weight effects in Russian. *Russian Linguistics*, *36*(3), 251–270.

Köhler, R. (1999). Syntactic structures: Properties and interrelations. *Journal of Quantitative Linguistics*, *6*(1), 46–57.

Kondo, T. & Yamashita, H. (2011). Why speakers produce scrambled sentences: an analysis of a spoken language corpus in Japanese. In H. Yamashita H Hirose & J. L. Packard (Eds.), *Processing and Producing Head-Final Structures* (pp. 195–217). New York, NY: Springer.

Konieczny, L. (2000). Locality and parsing complexity. *Journal of Psycholinguistic Research*, *29*(6), 627–645.

Konieczny, L. & Döring, P. (2003). Anticipation of clause-final heads: Evidence from eye-tracking and SRNs. In *Proceedings of iccs/ascs* (pp. 13–17).

Konopka, A. E. & Brown-Schmidt, S. (2014). Message encoding. In V. Ferreira, M. Goldrick, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 3–20). New York: Oxford University Press.

Kuno, S. (1974). The position of relative clauses and conjunctions. *Linguistic Inquiry*, *5*(1), 117–136.

Kurumada, C. & Jaeger, T. F. (2015). Communicative efficiency in language production: Optional case-marking in Japanese. *Journal of Memory and Language*, *83*, 152–178.

Kutas, M. & Hillyard, S. A. (1984). *Nature*, *307*, 161–163.

Kuznetsova, A., Brockhoff, P. B., & B.Christensen, R. H. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26.

Laka, I. (1994). *On the Syntax of Negation, Outstanding Dissertations in Linguistics*. A Garland Series.

Lambrecht, K. (1994). *Information structure and sentence form: Topic, Focus, and the Mental Representation of Discourse Referentes*. Cambridge, UK: Cambridge University Press.

Landa, A. (1995). *Conditions on null objects in Basque Spanish and their relation to leísmo and clitic doubling* (Unpublised doctoral dissertation, Los Angeles, CA: University of Southern California).

Lee, H. (2015). Usage probability and subject–object asymmetries in Korean case ellipsis: Experiments with subject case ellipsis. *Journal of Linguistics*, *52*, 70–110.

Lehman, W. P. (1973). A structural principle of language and its implications. *Language*, *49*, 47–66.

Levelt, W. J. M. & Maassen, B. (1981). Lexical search and order of mention in sentence production. In K. W & W. J. M. Levelt (Eds.), *Crossing the Boundaries in Linguistics* (pp. 221–252). Dordrecht, NL: D. Reidel.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177.

Levy, R., Fedorenko, E., & Gibson, E. (2013). The syntactic complexity of Russian relative clauses. *Journal of Memory and Language*, *69*(4), 461–495.

Levy, R. & Keller, F. (2013). Expectation and locality effects in German verb-final structures. *Journal of Memory and Language*, *68*(2), 199–222.

Lewis, R. L. (1996). Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research*, *25*(1), 93–115.

Lewis, R. L. & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419.

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447–454.

Linzen, T. & Oseki, Y. (2015). The reliability of acceptability judgments across languages. Manuscript.

Liu, H. (2008). Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science*, *9*(2), 159–191.

Liu, H. & Xu, C. (2012). Quantitative typological analysis of Romance languages. *Poznań Studies in Contemporary Linguistics*, *48*, 171–193.

Liu, H., Xu, C., & Liang, J. (2017). Dependency distance: A new perspective on syntactic patterns in natural languages. *Physics of Life Reviews*, *21*, 171–193.

Lohse, B., Hawkins, J. A., & Wasow, T. (2004). Domain minimization in English verb-particle constructions. *Language*, *80*(2), 238–261.

Lupyan, G. & Dale, R. (2010). Language structure is partly determined by social structure. *PloS one*, *5*(1), e8559.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, *9*(2), 157–201.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*, 226.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24*(1), 56–98.

MacDonald, M. C., Montag, J. L., & Gennari, S. P. (2016). Are there really syntactic complexity effects in sentence production? A reply to Scontras et al. (2015). *Cognitive Science*, *40*(2), 513–8.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676–703.

MacWhinney, B. (1987). The Competion Model. In K. W & W. J. M. Levelt (Eds.), *Mechanisms of Language Acquisition* (pp. 249–308). Hillsdale, NJ: Lawrence Erlbaum.

MacWhinney, B. & Bates, E. (1989). The Cross-linguistic study of sentence processing. *The American Journal of Psychology*, *105*(3), 484–492.

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 127–150.

Mahowald, K., Fedorenko, E., Piantadosi, S. T., & Gibson, E. (2013). Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition*, *126*(2), 313–318.

Maia, J. (2015). Gazteek atzerkarga handiagoa nahi dute euskaraz? *Uztaro*, *95*, 51–71.

Martin, A. E. & Freedman, M. L. (2001). Short-term retention of lexical-semantic representation: Implications for speech production. *Memory*, *9*(4-6), 261–280.

Martin, A. E., Miller, M., & Vu, H. (2004). Lexical-semantic retention and speech production: Further evidence from normal and brain damaged participants for a phrasal scope of planning. *Cognitive Neuropsychology*, *21*, 625–644.

Maurits, L., Navarro, D., & Perfors, A. (2010). Why are some word orders more common than others? A uniform information density account. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 23* (pp. 1585–1593). Curran Associates, Inc.

McDonald, J. L., Bock, K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological and metrical determinants of serial position. *Cognitive Psychology*, *25*(2), 188–230.

McElree, B. (2000). Sentence comprehension Is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, *29*(2), 111–123.

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, *48*(1), 67–91.

Merlo, P. (2015). Evaluation of two-level dependency representations of argument structure in long-distance dependencies. In *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)* (pp. 221–230).

Miller, G. A. & Chomsky, N. (1963). Finitary models of language users. *Advances in Neural Information Processing Systems*, *23*, 1585–1593.

Myachykov, A., Scheepers, C., Garrod, S., Thompson, D., & Fedorova, O. (2013). Syntactic flexibility and competition in sentence production: The case of English and Russian. *The Quarterly Journal of Experimental Psychology*, *66*(8), 1601–1619.

Myachykov, A., Thompson, D., Scheepers, C., & Garrod, S. (2011). Visual attention and structural choice in sentence production across languages. *Language and Linguistics Compass*, *5*(2), 95–107.

Nairne, J. . S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, *53*(1), 53–81.

Nakatani, K. & Gibson, E. (2010). An on-line study of Japanese nesting complexity. *Cognitive Science*, *34*(1), 94–112.

Neurobehavioral Systems Inc. (2016). Presentation. http : www . neurobs . com. Computer Sofware. Berkeley, CA.

Newport, E. L. (1999). Reduced input in the acquisition of signed languages: Contributions to the study of creolization. In M. Degraff (Ed.), *Creolization, Diachrony, and Language Acquisition* (pp. 161–78). Cambridge, MA: MIT Press.

Nicenboim, B., Logačev, P., Gattei, C., & Vasishth, S. (2016). When high-capacity readers slow down and low-capacity readers speed up: Working memory and locality effects. *Frontiers in Psychology*, *7*, 280.

Nicenboim, B., Vasishth, S., Gattei, C., Sigman, M., & Kliegl, R. (2015). Working memory differences in long-distance dependency resolution. *Frontiers in Psychology*, *6*, 312.

Norcliffe, E., Harris, A. C., & Jaeger, T. F. (2010). Cross-linguistic psycholinguistics and its critical role in theory development: early beginnings and recent advances. *Language, Cognition and Neuroscience*, *30*(9), 1009–1032.

Norman, S., Kemper, S., Kynette, D., Cheung, H., & Anagnopoulos, C. (1991). Syntactic complexity and adults' running memory span. *Journal of gerontology*, *46*(6), P346–P351.

Open Science Collaboration and others. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716.

Ortiz de Urbina, J. (1989). *Parameters in the grammar of Basque: A GB approach to Basque syntax*. Foris Pubns USA.

Pablos, L. (2011). Rich agreement in Basque: Evidence for pre-verbal structure building. In H. Yamashita, Y. Hirose, & J. L. Packard (Eds.), *Processing and producing head-final structures* (pp. 3–21). New York, NY: Springer.

Park, Y. A. & Levy, R. (2009). Minimal-length linearizations for mildly context-sensitive dependency trees. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 335–343). Association for Computational Linguistics.

Pastor, L. & Laka, I. (2013). Processing facilitation strategies in OV and VO languages: A corpus study. *Open Journal of Modern Linguistics*, *3*(3), 252–258.

Pearlmutter, N. J. & MacDonald, M. C. (1995). Individual differences and probabilistic constraints in syntactic ambiguity resolution. *Journal of Memory and Language*, *34*(4), 521–542.

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, *122*(3), 280–91.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, *108*(9), 3526–3529.

Prat-Sala, M. & Branigan, H. P. (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. *Journal of Memory and Language*, *42*(2), 168–182.

Primus, B. (1999). *Case and thematic roles: Ergative, accusative and active*. Tübingen: Niemeyer.

Prince, E. F. (1995). On the limits of syntax, with reference to left-dislocation and topicalization. In P. W. Culicover & L. McNally (Eds.), *The limits of syntax* (pp. 281–302). Academic Press.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A Grammar of contemporary English*. London: Longman.

R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Rajkumar, R., van Schijndel, M., White, M., & Schuler, W. (2016). Investigating locality effects and surprisal in written English syntactic choice phenomena. *Cognition*, *155*, 204–232.

Regier, T., Kemp, C., & Kay, P. (2015). Word meanings across languages support efficient communication. In B. MacWhinney & W. O'Grady (Eds.), *The Handbook of Language Emergence* (pp. 237–263). Wiley-Blackwell.

Rickford, J., Wasow, T., Mendoza-Denton, N., & Espinoza, J. (1995). Syntactic variation and change in progress: Loss of verbal coda in topic restricting *as far as* constructions. *Language*, *71*, 101–131.

Ros, I., Santesteban, M., Fukumura, K., & Laka, I. (2015). Aiming at shorter dependencies: the role of agreement morphology. *Language, Cognition and Neuroscience*, *30*(9), 1156–1174.

Ross, J. R. (1967). *Constraints on variables in syntax* (Doctoral dissertation, PhD thesis. Cambridge, MA: Massachusetts Institute of Technology).

Safavi, M. S., Husain, S., & Vasishth, S. (2016). Dependency resolution difficulty increases with distance in Persian separable complex predicates: Evidence for expectation and memory-based accounts. *Frontiers in Psychology*, *7*, 403.

Saito, M. (2004). Japanese scrambling in a comparative perspective. In D. Adger, D. de Cat, & G. Tsoulas (Eds.), *Peripheries: Syntactic Edges and their Effects* (pp. 143–163). Dordrecht, NL: Kluwer Academic Publishers.

Sasaki, Y. & MacWhinney, B. (2006). The Competition Model. In M. Nakayama & P. Li (Eds.), *The Handbook of East Asian Psycholinguistics* (pp. 307–314). Cambridge, UK: Cambridge University Press.

Schweppe, J., Barth, S., Ketzer-Noltge, A., & Rummer, R. (2015). Does verbatim sentence recall underestimate the language competence of near-native speakers? *Front Psychol*, *6*, 63.

Scontras, G., Badecker, W., & Fedorenko, E. (2017). Syntactic Complexity Effects in Sentence Production: A Reply to MacDonald, Montag, and Gennari (2016). *Cognitive science*, *41*(8), 2280–2287.

Scontras, G., Badecker, W., Shank, L., Lim, E., & Fedorenko, E. (2015). Syntactic complexity effects in sentence production. *Cognitive Science*, *39*(3), 559–83.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423.

Siewierska, A. (1993). Syntactic weight vs. pragmatic factors and word order variation in Polish. *Journal of Linguistics*, *29*, 233–265.

Silva-Corvalán, C. (1984). Semantic and pragmatic factors in syntactic change. In J. Fisiak (Ed.), *Historical syntax* (pp. 555–573). Berlin: Mouton.

Singleton, J. L. & Newport, E. L. (2004). When learners surpass their models: The acquisition of American Sign Language from impoverished input. *Cognitive Psychology*, *49*, 370–407.

Sleator, D. D. & Temperley, D. (1991). Parsing English with a Link Grammar. Parsing English with a link grammar. Technical report CMY-CS-91-196. Pittsburgh, PA: Carnegie Mellon University, School of Computer Science.

Slevc, L. R. (2011). Saying what's on your mind: Working memory effects on sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1503.

Slevc, L. R. & Martin, R. C. (2016). Syntactic agreement attraction reflects working memory processes. *Journal of Cognitive Psychology*, *28*(7), 773–790.

Slobin, D. I. (2001). Cognitive prerequisites for the development of grammar. In D. I. Slobin & C. A. Ferguson (Eds.), *Studies of Child Language Development* (Vol. 12, pp. 175–208). New York, NY: Holt, Rinehart, and Winston.

Smith, M. & Wheeldon, L. (2004). Horizontal information flow in spoken sentence production. *Journal of Experimental Psychology: Learning Memory and Cognition*, *30*, 675–686.

Smith, N. J. & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128*(3), 302–19.

Sprouse, J. & Almeida, D. (2012). Assessing the reliability of textbook data in syntax: Adger's Core Syntax. *Journal of Linguistics*, *48*(3), 609–652.

Stallings, L. M. & MacDonald, M. C. (2011). It's not just the "heavy NP": Relative phrase length modulates the production of heavy-NP shift. *Journal of Psycholinguistics*, *40*(3), 177–87.

Stallings, L. M., MacDonald, M. C., & O'Seaghdha, P. G. (1998). Phrasal ordering constraints in sentence production: Phrase length and verb disposition in heavy-NP shift. *Journal of Memory and Language*, *39*(3), 392–417.

Staub, A. (2010). Eye movements and processing difficulty in object relative clausesr. *Cognition*, *116*, 71–86.

Swets, B., Desmet, T., Hambrik, D. Z., & Ferrerira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*, *136*(1), 64–81.

Swets, B., Jacovina, M. E., & Gerrig, R. J. (2014). Individual differences in the scope of speech planning: evidence from eye-movements. *Language and Cognition*, *6*(1), 12–44.

Szmrecsányi, B. (1999). On operationalizing syntactic complexity. *Journées Internationals d'Analyse statistique des Données Textualles*, *73*(3), 205–246.

Tanaka, M. N., Branigan, H. P., & Pickering, M. J. (2011). The Production of Head-initial and Head-final languages. In H. Yamashita, Y. Hirose, & J. L. Packard (Eds.), *Processing and producing head-final structures* (pp. 113–129). New York, NY: Springer.

Tanenhaus, M. K. (2004). On-line sentence processing: past, present and future. In M. Carreiras & C. J. Clifton (Eds.), *On-line sentence processing: ERPs, Eye Movements and Beyond* (pp. 371–392). Brighton, UK.: Psychology Press.

Tanenhaus, M. K. (2013). All P's or mixed vegetables? *Frontiers in Psychology*, *4*, 234.

Tanenhaus, M. K. & Trueswell, J. C. (1995). Sentence comprehension. In J. L. Miller & P. D. Eimas (Eds.), *Handbook of Perception and Cognition Vol. 11, Speech, Language and Communication* (pp. 371–392). San Diego, CA: Academic Press.

Taulé, M., Martı, M. A., & Recasens, M. (2008). AnCora: Multilevel Annotated Corpora for Catalan and Spanish. In *Lrec*.

Tayor, W. L. (2015). "Cloze procedure": A new tool for measuring readibility. *Journalism Quaterly*, *30*, 415–433.

Temperley, D. (2007). Minimization of dependency length in written English. *Cognition*, *105*(2), 300–33.

Temperley, D. (2008). Dependency-length minimization in natural and artificial languages. *Journal of Quantitative Linguistics*, *15*(3), 256–282.

Temperley, D. & Gidea, D. (2010). Do grammars minimize dependency length? *Cognitive Science*, *34*, 286–310.

Temperley, D. & Gidea, D. (2018). Minimizing syntactic dependency lengths: Typological/Cogntive Universal? *Annuel Review of Linguistics*, *4*, 1–14.

Tesnière, L. (1959). *Eléments de syntaxe structurale*. Klincksieck.

Tily, H. J. (2010). *The role of processing complexity in word order variation and change* (Unpublished doctoral dissertation, PhD thesis. Stanford University).

Tippets, I. (2011). Differential object marking: Quantitative evidence for underlying hierarchical constraints across Spanish dialects. In *13th Hispanic Linguistics Symposium* (pp. 107–117).

Tomlin, R. S. (1986). *Basic Word Order: Functional Principles*. London/Wolfeboro, N.H.: Croom Helm.

Torrego, E. (1998). *The dependencies of objects*. Cambridge, MA: MIT Press.

Trask, R. L. (1997). *The history of Basque*. London: Routledge.

Trueswell, J. C. & Tanenhaus, M. K. (1993). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. C & F. L (Eds.), *Perspectives on sentence processing* (pp. 155–179). NJ: Lawrence Erlbaum & Associates.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(3), 528–553.

Turner, M. L. & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154.

Ueno, M. & Polinsky, M. (2009). Does headedness affect processing? A new look at the VO–OV contrast. *Journal of Linguistics*, *45*(3), 675–710.

Unsworth, N., Heitz, P. R., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505.

Uszkoreit, H. (1987). *Word order and constituent structure in German*. Standford, CA: CSLI.

Uszkoreit, H., Brants, T., Duchier, D., Krenn, B., Konieczny, L., Oepen, S., & Skut, W. (1998). Studien zur performanzorientierten Linguistik: Aspekte der Relativsatzextraposition in Deutschen. [Studies in performance-oriented linguistics: Aspects of German relative clause extraposition]. *Kognitionswissenschaft*, *7*(3), 129–133.

Van Dyke, J. A., Clinton, L. J., & Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition*, *131*, 373–403.

Van Dyke, J. A. & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*(3), 285–316.

Van Dyke, J. A. & McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*, *65*(3), 247–263.

Vasishth, S. & Drenhaus, H. (2011). Locality in German. *Dialogue and Discourse*, *2*(1), 59–82.

Vasishth, S. & Lewis, R. L. (2006). Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language*, 767–794.

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 423–440.

Wang, Y. & Liu, H. (2017). The effects of genre on dependency distance and dependency direction. *Language Sciences*, *59*, 135–147.

Wanner, H. E. & Maratsos, M. P. (1978). An ATN approach to comprehension. In M. A. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic Theory and Psychological Reality*. MIT Press.

Wasow, T. (1997a). End weight from the speaker's perspective. *Journal of Psycholinguistics Research*, *26*(3), 347–361.

Wasow, T. (1997b). Remarks on grammatical weight. *Language Variation and Vhange*, *9*(01), 81–105.

Wasow, T. (2002). *Postverbal behavior*. Stanford, CA: CSLI.

Wasow, T. (2013). The appeal of the PDC program. *Frontiers in Psychology*, *4*, 236.

Waters, G. S. & Caplan, D. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review*, *103*, 761–772.

Weissenrieder, M. (1990). Variable uses of the direct-object marker *a*. *Hispania*, *73*, 223–231.

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, *58*(2), 250–71.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wiechmann, D. & Lohmann, A. (2013). Domain minimization and beyond: Modelling prepositional phrase ordering. *Language Variation and Change*, *25*(1), 65–88.

Yamashita, H. & Chang, F. (2001). "Long before short" preference in the production of a head-final language. *Cognition*, *81*(2), B45–B55.

Yetano, I., Duñabeitia, J. A., & Laka, I. (2012, June). Agent-initial sentence processing preference: A visual-word eye-movement experiment. In *Poster presented at the 7th International Morphological Processing Conference*. Donostia-San Sebastián, Spain.

Yngve, V. (1960). A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society*, *104*(5), 444–446.

Zec, D. & Inkelas, S. (1990). Prosodically constrained syntax. In D. Zec & S. Inkelas (Eds.), *The Phonology-Syntax Connection* (pp. 365–378). Standford, CA: Center for the Study of Language and Informations.

Zeman, D., Marecek, D., Popel, M., Ramasamy, L., Stepánek, J., Zabokrtský, Z., & Hajic, J. (2012). HamleDT: To Parse or Not to Parse? In *LREC* (pp. 2735–2741).

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley.

Zubimendi, J. R. & Esnal, P. (1993). *Idazkera-liburua*. Vitoria-Gasteiz: Eusko Jaurlaritza-Kultura Saila.

Zubizarreta, A. M. (1994). Word Order in Spanish the Nature of Nominative Case. In K. Johnson & I. Roberts (Eds.), *Beyond Principles and Parameters. Essays in Memory of O. Jaeggli* (pp. 223–250). Dordrecht, NL: Kluwer Academic Press.