FACULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES

Listening to Native and Non-native Speech: Prediction and Adaptation

Aster Dijkgraaf

Supervisor: Prof. Dr. Wouter Duyck
Co-supervisor: Prof. Dr. Robert Hartsuiker

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Doctor of Psychology

Academic year 2017–2018

GHENT
UNIVERSITY

Listening to Native and Non-native Speech: Prediction and Adaptation

Aster Dijkgraaf

Supervisor: Prof. Dr. Wouter Duyck
Co-supervisor: Prof. Dr. Robert Hartsuiker

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Doctor of Psychology

Academic year 2017–2018

CONTENTS

ACKNOWLEDGEMENTS **5**

CHAPTER 1 INTRODUCTION **9**

Listening in Context 10
Prediction in Native Language Comprehension: Evidence, Mechanisms, Modulators 12
Listening in the Non-native Language: What is Different? 17
Prediction in Non-native Language Comprehension 21
Speech Production after Speech Perception in the Non-native Language 27
Current Dissertation 28
References 31

CHAPTER 2 PREDICTING UPCOMING INFORMATION IN NATIVE-LANGUAGE AND NON-NATIVE-LANGUAGE AUDITORY WORD RECOGNITION **45**

Introduction 46
Methods 53
Results 59
Discussion and Conclusion 62
References 67

CHAPTER 3 PREDICTION AND INTEGRATION OF SEMANTICS DURING L2 AND L1 LISTENING **73**

Introduction 74
Method 80
Results 90
Discussion 97
References 103

CHAPTER 4 PREDICTION OF SEMANTICS IN NATIVE AND NON-NATIVE SPEECH COMPREHENSION: THE ROLE OF COGNITIVE LOAD AND PROCESSING SPEED **113**

Introduction 114
Experiment 1 120
Experiment 2 130
General Discussion 138
Conclusion 143
References 144

CHAPTER 5 IS THERE ADAPTATION OF SPEECH PRODUCTION AFTER SPEECH PERCEPTION IN BILINGUAL INTERACTION? **151**

Introduction 152
Method 157
Results 167
Discussion 173
References 177

CHAPTER 6 GENERAL DISCUSSION **181**

Empirical Findings 181
Theoretical Implications 186
Limitations and Future Research 194

Conclusion 197
References 198

CHAPTER 7 ENGLISH SUMMARY **203**

References 207

CHAPTER 8 NEDERLANDSE SAMENVATTING **211**

References 216

APPENDICES CHAPTER 2 **221**

Appendix 2A 222

APPENDICES CHAPTER 3 **223**

Appendix 3A 223
Appendix 3B 236
Appendix 3C 238
Appendix 3D 240

APPENDICES CHAPTER 4 **244**

Appendix 4A 244
Appendix 4B 247

APPENDICES CHAPTER 5 **263**

Appendix 5A 263
Appendix 5B 275
Appendix 5C 276

DATA STORAGE FACT SHEET CHAPTER 2 **281**

DATA STORAGE FACT SHEET CHAPTER 3 **283**

DATA STORAGE FACT SHEET CHAPTER 4 **285**

DATA STORAGE FACT SHEET CHAPTER 5 **287**

ACKNOWLEDGEMENTS

This dissertation would not have come about without the help of many wonderful people.

First, I would like to express my gratitude to these four researchers whom I have had the pleasure of meeting during my master's programme: Frank Wijnen, Gaetano Fiorin, Doug Davidson and Jos van Berkum. You have introduced me to many different aspects of scientific research, but more importantly, you have sparked my enthusiasm for pursuing a PhD in psycholinguistics time after time. Jos van Berkum, I often reread some of the literature you introduced me to in your course on language and emotion. That helped me to see the bigger picture and kept me excited to work on prediction in language comprehension.

I would like to thank my doctoral supervisors Wouter Duyck and Rob Hartsuiker for their excellent guidance throughout the past years. Wouter, thank you for the opportunity to start this project and for trusting me to plan my work. I very much appreciated your advice and your sharp interpretation of scientific results. Rob, I admire your extensive knowledge of the academic literature and the way in which you convey your enthusiasm for scientific research. I learned a lot from your comments on my writing and research plans.

I am also indebted to the other members of my doctoral guidance committee, Marc Brysbaert, Martin Valcke, Andrea Weber, Evelyne Lagrou and Evy Woumans for their valuable input and advice throughout the project.

Sarah Bernolet, Alix Priano and Isa Brysbaert have spent many hours recording sentences for my experiments. Thank you very much for your time and patience. Here, I would also like to mention Michael Stevens, Eva Marie Koch and Kyla Robinson who all contributed to the construction of the stimuli and recordings for my research, thank you.

Many thanks to Denis Drieghe (University of Southampton) for the opportunity to test a group of monolingual participants in Southampton and for his kind hospitality during my stay.

Eva van Assche, thank you for your advice and for your collaboration on various projects.

And then there are my PhD colleagues:

I must begin with one colleague without whom completing this PhD project would not have been possible: Eowyn Van de Putte. Thank you for all the wonderful conversations, for many fun dinners, for the walk through the sea in Oostend, for pimientos padrones for dessert and of course for always having a room available for me in Ghent when I needed one.

Uschi, Hanne, en Hanna, thank you for many memorable evenings, for our fun and interesting conversations and for your support.

Raquel, my neighbour at the department, thank you for many chats that somehow always lifted my spirits.

Heleen, LEMMA project's binding factor. I admire your communicative skills, your kindness and your inspiring enthusiasm. I very much appreciated our conversations about science and about completing a PhD.

Wouter Broos, thank you for your hard work on our joint project. I have fond memories of eating fries at the office and of having skype conversations about *blobs*, *babes* and *landscapes*.

Nicolas, my office mate, thank you for the nice chats and for answering many questions about R. Your drive and efficiency were very motivating.

Toru, my fellow LEMMA from day one, thank you for sharing your comments and ideas and for always patiently answering my statistics questions.

Ellen and Marcelo, thank you for sharing your research and ideas on many LEMMA meetings and for the occasional drink together.

I will switch languages now to address my Dutch friends and family, as expressing my gratitude to them feels more natural in my L1. Ten slotte wil ik graag mijn vrienden, familie en schoonfamilie bedanken voor alle steun, interesse, en geduld met al mijn twijfels. Een paar mensen wil ik hier graag nog speciaal vermelden.

Giulia, you were there from my very first steps into academia. You share the love I have for language, but as you soon found out, not for linguistics. I am grateful that when we no longer shared our academic studies we continued to share coffees, which turned into wines and many wonderful dinners. Thank you for your love, support, and for all the amazing memories.

Anna, Adinda en Maaike, jullie zijn al bijna twintig jaar mijn steun en toeverlaat. De afgelopen jaren waren daarop geen uitzondering, enorm bedankt daarvoor.

Jeroen en Malou, de afgelopen jaren hebben we lief en leed gedeeld. Ik waardeer onze vriendschap zeer.

Mijn zus en broertje, Hannah en Lucas, heel erg bedankt voor jullie interesse in mijn onderzoek en voor de vele gezellige, ontspannende momenten. Oma Atie, hartelijk bedankt voor je steun. Je hebt een grotere rol gespeeld bij het tot stand komen van mijn proefschrift dan je weet.

Irma en Jan, mijn lieve schoonouders en geweldige opa en oma voor mijn kinderen, heel erg bedankt voor al jullie steun de afgelopen jaren.

Papa en mama, ik kan jullie niet genoeg bedanken voor jullie steun tijdens mijn opleiding en promotietraject. Jullie grote enthousiasme voor taal en wetenschap ligt ook in mij verankerd en heeft zonder twijfel geleid tot de keuze voor het mooie vakgebied taalwetenschap. Jullie liefde heeft me de afgelopen jaren door veel moeilijke momenten geholpen. Bovendien was jullie hulp en die van Jan en Irma met de kinderen de afgelopen jaren onmisbaar en was ons huis al lang ingestort zonder het harde werk van papa en Jan.

Mijn allerliefste meiden, Lola en Maite. Door jullie aanwezigheid geniet ik iedere dag. En dan nog Joost. Jouw vertrouwen en optimisme gaven mij de moed en de wil dit proefschrift af te maken. Dank voor je eindeloze geduld. Je maakt mij ongelofelijk gelukkig.

CHAPTER 1
INTRODUCTION

Bilinguals have been defined as those people who use two or more languages in their everyday lives (Grosjean, 2010). This definition entails that, contrary to popular belief, bilinguals do not necessarily have native-like mastery of two or more languages. Bilinguals range from people that learn and use two or more languages from birth, to people that learned a second language as adult and use it only in a particular context. Bilingualism is therefore not a rare phenomenon. It has been estimated that more than half of the world's population has knowledge of two or more languages (Grosjean, 2010). In Belgium, where this dissertation came about, 58% of adults participating in the European Union's Adult Education Survey had knowledge of two or more languages in 2011 ("Foreign language skills statistics - Statistics Explained," 2015).

Thus, monolinguals are not the standard or default language users, and it is therefore important that theories and models of language processing are not only designed for the monolingual case, but that they are generalized to apply to native (L1) and non-native (L2) processing in bilinguals as well. This is not an easy task: L1 and L2 processing by bilinguals differs from the monolingual case in many ways, and one bilingual is not the other. Bilinguals differ wildly from each other (and from monolinguals) in terms of proficiency, age of acquisition, cultural background, language use, frequency of language switches, culture, socio-economic status and many more. It is therefore important that researchers keep in mind that findings that apply to one group of bilinguals may not necessarily apply to another, and that monolinguals may differ from bilinguals in relevant aspects other than language experience.

A lot of research has focussed on exploring the differences between monolingual and bilingual language processing, with one of the major research lines focusing on cross-linguistic interference due to parallel language activation in bilinguals (e.g., Dijkstra, Grainger, & van Heuven, 1999; Duyck, Assche, Drieghe, & Hartsuiker, 2007). In this thesis, we focus mainly on differences in spoken language processing between L1 and L2 within the same bilingual individuals. This research is aimed at advancing our understanding of mechanisms involved in human speech comprehension in general. And, at a more practical level, this type of research helps us understand disadvantages in L2 comprehension in (increasingly widespread) bilingual societies. This is important, as bilinguals do not only use L2 in informal conversation, but the L2 is also increasingly used in formal settings such as in higher education and in professional contexts. One large advantage of the within-participants approach employed here, is that this way, any difference in knowledge, life experience and background between L1 and L2 comprehenders is eliminated, except for the differences in language experience.

In the next sections, we will first introduce L1 spoken language comprehension and we will zoom in on the main topic studied in this dissertation: prediction of upcoming information during language comprehension. Then, we will continue to discuss differences between L1 and L2 comprehension and how these differences may interfere with prediction during comprehension. Next, we present an overview of studies focusing on prediction in the L1 and the L2. Further, we discuss how top-down effects in speech comprehension can also affect speech production and why this mechanism may fail in the L2. This introduction section is concluded with an overview of the goals and the chapters in the current dissertation.

LISTENING IN CONTEXT

Understanding speech in our L1 is a skill that is often taken for granted. We usually understand each other without any effort and we learned to do this, even before formal instruction, at a very early age. However, understanding speech entails many complicated processes. For a start, we need to segment the continuous speech signal into sounds and words. This is complex because unlike written language, speech unfolds over time. Segments (words, phonemes, sentences) overlap and silences are often not an informative cue about word or sentence boundaries. Besides segmentation, phonemes and words need to be recognized and combined into larger units (constituents, sentences and discourse) based on syntactic rules, so that meaning can be derived them. A listener is further challenged by not being able to go back to verify a previous segment like a reader can. And on top of that, speech is fast (about 150-190 words per minute) and highly variable due to characteristics of the speaker (such as voice pitch, age, gender, dialect, mood, having a cold), but also due to other exterior factors such as background noise. Finally, speech is often highly ambiguous. The same sequence of phonemes can be segmented into different words (e.g. *ice cream* - *I scream*), and some phrases can initially be parsed in more than one way (*The man who whistles tunes pianos*). Pragmatics and prosody may also affect meaning, for example by indicating intended irony. Thus, when we consider all the factors implicated in speech perception, the ease with which we understand spoken language is rather extraordinary.

To cope with inherently fast, noisy and ambiguous spoken language, the comprehension system uses each bit of input incrementally; that is, information is integrated and interpreted as soon as it becomes available. This happens at all levels. We do not wait with word recognition before the entire word is heard, but we already activate potential word candidates upon hearing the first sounds of a word. Syntactic structure is build up as soon as a constituent comes in and not after hearing an entire sentence. For instance, when hearing *the man who whistles tunes pianos*, *tunes* is initially integrated as direct object. Only upon hearing the disambiguating word *pianos,* the sentence is parsed in the correct way, with *tunes* as verb. The processing difficulty associated with the disambiguating word in such a (temporarily ambiguous) sentence is called the garden-path-effect. We also start constructing sentence meaning before the entire sentence is heard. This way, sentence context can ease the semantic integration of an incoming word in the sentence. For instance, the word *butter* is processed more easily than the word *socks* when it follows the sentence *He spread his warm bread with ...* (Kutas & Hillyard, 1980). There is ample behavioral (e.g., Boland, Tanenhaus, Garnsey, & Carlson, 1995; Sedivy, K Tanenhaus, Chambers, & Carlson, 1999; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) and neural (e.g., Kutas & Hillyard, 1984; van Berkum, Brown, & Hagoort, 1999) evidence for incremental interpretation in language processing.

One major source of evidence for such incrementality is provided by visual world eye-tracking studies. In these types of studies, participants typically listen to spoken language and look at objects depicted on a screen while their eye-movements are measured. In a study by Sedivy et al. (1999), for example, participants viewed displays with two objects that only differed in one property (such as colour, e.g., a pink and a yellow comb), one different object that shared a property with one member of the object pair (e.g. a yellow bowl) and an unrelated object (e.g. a metal knife). Upon hearing instructions such as *"Touch the pink comb. Now touch the yellow comb/bowl"*, participants were much faster to look at the contrast referent (comb) than at the non-contrast referent. This experiment shows that people immediately direct eye-movements to objects compatible with the visual and linguistic input. This suggests that comprehenders construct sentence meaning incrementally, as each new word comes in. Importantly, in the past two decades or so, more and more studies found evidence showing that, on top of incremental interpretation, people actually use context information to make predictions about upcoming input during language comprehension (Altmann & Kamide, 1999; DeLong, Urbach, & Kutas, 2005; Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005).

PREDICTION IN NATIVE LANGUAGE COMPREHENSION: EVIDENCE, MECHANISMS, MODULATORS

"So in all, our brains are not just proactive because it pays to anticipate upcoming events in a complex dynamic world. It is also because the input would otherwise simply be too difficult to deal with efficiently." (Van Berkum, 2010, p. 5)

Prediction of upcoming information is considered to be a key principle in many subfields of human cognition (Bar, 2007, 2009), and human language processing is no exception (Kuperberg & Jaeger, 2016; Kutas, DeLong, & Smith, 2011; Van Berkum, 2010; van Berkum, 2013). Although the extent to which language users (need to) engage in prediction is subject to debate (see for instance, Huettig & Mani, 2016; Kuperberg & Jaeger, 2016), accumulating evidence shows that comprehenders can use all kinds of linguistic (e.g., semantic, syntactic, prosodic, phonological, discourse) and non-linguistic information (such as visual context and background knowledge) to predict upcoming linguistic information during language comprehension at all levels. Predictive processing can be beneficial for comprehenders in several ways. For instance, predictive processing can help comprehenders to deal with fast, noisy and ambiguous input, it can give the comprehender a head start on future material (as long as predictions are correct), and it can help to determine when it is time to start an overt response in dialogue (e.g., Kutas et al., 2011; Van Berkum, 2010).

Evidence from Behavioral Studies

Behavioral research on prediction during speech comprehension mainly used the visual world paradigm (see Huettig, Rommers, & Meyer, 2011, for a review), in which participants' eye-movements are measured while they look at a display and listen to an utterance (Cooper, 1974). In a seminal study by Altmann and Kamide (1999), participants viewed semi-realistic visual scenes and listened to sentences containing for example a boy, a cake, and some toys. Participants would hear sentences with either a constraining or a neutral verb given the visual context: *The boy will eat/move the cake* where there was only one edible object in the display (*the cake)* and all objects could be moved. Participants fixated the cake earlier when hearing the constraining verb *eat* than when hearing the neutral verb *move*, before information of the final noun (*cake*) could affect processing. This was taken as evidence for semantic pre-activation of the word *cake*.

This semantic prediction effect has been replicated many times with similar paradigms. There is now evidence from visual world studies that people integrate information from the visual context and world-knowledge with verb semantics (Altmann & Kamide, 1999), information from the grammatical subject and the verb (Kamide, Altmann, & Haywood, 2003), verb tense (Altmann &

Kamide, 2007), verb transitivity (Boland, 2005), or case-marking information (Kamide et al., 2003) and use this integrated information to generate predictions about upcoming speech input. Knoeferle et al. (2005) extended these findings by showing that people rapidly integrate visual and linguistic information without much use of world-knowledge stored in long-term memory when generating predictions. Participants were shown unusual visual scenes (a princess washing a pirate while a fencer is painting the princess). When hearing sentences describing the scene such as *Die Prinzessin wäscht offensichtlich den Pirat* (The princess$_{nom/acc}$ washes apparently the pirate$_{acc}$), participants integrated information from the verb and from the visual scene to disambiguate the case of the first referent and anticipate the upcoming referent.

Evidence from ERP Studies

Another large body of evidence for predictive pre-activation comes from event-related potential studies (e.g., DeLong et al., 2005; Otten & Van Berkum, 2008; Van Berkum et al., 2005, 2005; Wicha, Bates, Moreno, & Kutas, 2003; Wicha, Moreno, & Kutas, 2004). One advantage of this type of research compared with visual world paradigm studies is that no visual context is needed to accompany the auditory input of interest. The cues used for prediction are only linguistic in nature and prediction effects are found before the target for prediction is encountered in the input. Therefore, results from these experiments are incompatible with facilitated integration accounts of context effects, in which it is assumed that prior context can facilitate processing of incoming information but that no pre-activation takes place. In these studies, participants are usually exposed to constraining sentences that support prediction of a specific noun. To test whether the noun is predicted before the actual onset of the noun in the stimulus sentence, a prenominal article or adjective is manipulated to be congruent or incongruent with the predictable noun. A differential ERP response elicited by the congruent compared to the incongruent article or adjective is taken to indicate that the noun is pre-activated. For example, in an EEG experiment by Van Berkum, Brown, Zwitserlood, Kooijman and Hagoort (2005) participants heard a discourse in Dutch such *The burglar had no trouble locating the secret family safe. Of course, it was situated behind a ...* followed by either (1) (consistent) or (2) (inconsistent) below.

a big$_{-\varnothing neu}$ but unobtrusive painting$_{neu}$ (neuter gender, adjective has "zero" suffix)

a big$_{-e\ com}$ but unobtrusive bookcase$_{com}$. (common gender, adjective has –e suffix)

The N400 response elicited by the prenominal adjective *big* was larger in the prediction-consistent than in the prediction-inconsistent condition, indicating that the word *painting* and its gender-feature was anticipated. In a study with a similar paradigm by Delong et al., participants read sentences varying in constraint with expected or less expected article/noun pairings: '*The day was breezy so the boy went outside to fly … a kite$_{[EXPECTED]}$/an airplane $_{[UNEXPECTED]}$ in the park'*. The prediction-congruent article *a* elicited a smaller N400 effect than the prediction-incongruent article *an*. Interestingly, there was an inverse correlation between the N400 response elicited by the noun and its cloze probability and the same inverse correlation was found for the N400 response to the article. This finding was taken to indicate that participants predicted target noun semantics and phonological form in a graded fashion. It should be noted, however, that a recent large-scale multi-lab replication attempt of Delong et al. failed to find the same effect of word form prediction on the article ERP (Nieuwland et al., 2017).

Routes and Mechanisms

Predictive processing figures prominently in recent theories of language comprehension (Altmann & Mirković, 2009; MacDonald, 2013; Pickering & Garrod, 2013). But how does prediction come about? The literature often distinguishes at least two routes to prediction (Huettig, 2015; Kuperberg, 2007; Pickering & Gambi, 2018; Pickering & Garrod, 2013). The first route is based on low-level associative relationships. Pre-activation occurs due to lingering activation from lower level representations of prior context. This route is often associated with priming. Prediction via associative relationships is usually assumed to be relatively automatic, in that it occurs involuntarily and that it requires no or few cognitive resources. The second route to predictive pre-activation uses higher-level (message-level) information derived from the context to actively pre-activate information at lower levels (Huettig, 2015; Kuperberg, 2007; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018; Pickering & Garrod, 2013). This route is usually assumed to be more resource and time consuming and possibly strategic (e.g., Huettig & Janse, 2016; Pickering & Gambi, 2018).

Otten and Van Berkum (2008), used EEG to test directly whether effects of prediction based on higher level information could be distinguished from effects of priming (automatic spreading activation based on low-level associations) by contrasting neural responses to anomalous words in a discourse context which was either highly predictive for a specific word, or non-predictive but containing the same prime words as the predictable context (e.g. *Sylvie and Joanna really feel like dancing and flirting tonight. Therefore they go to a* ***stove*** *[disco] (...)* $_{[PREDICTIVE\ CONTEXT]}$ versus *After* *all the dancing Sylvie and Joanna really don't feel like flirting tonight. Therefore they go to a* ***stove*** *[disco](...) [NON-PREDICTIVE CONTEXT]*). Neural responses to targets differed for prime control stories and predictive stories, suggesting that comprehenders did not (or at least not only) pre-activate words due to simple priming mechanisms, but that they can use higher order, message level information to generate predictions. Other studies have provided converging evidence for prediction based on higher level information and (untargeted) prediction based on automatic spreading activation (Hintz, 2015; Kukona, Fang, Aicher, Chen, & Magnuson, 2011; Lau, Holcomb, & Kuperberg, 2013). For instance, Kukona et al. used a visual world paradigm to distinguish effects of active pre-activation and passive spreading of activation. Participants were exposed to sentences such as '*Toby arrests the crook'* And viewed displays including a likely subject for the verb (*policeman*) and a likely object (*crook*). Upon hearing the verb, participants anticipated verb-related agents and verb-related patients almost to the same extent (even though the agent role was already filled), suggesting simultaneous effects of passive priming and active prediction. When the sentences were presented in OVS order (*'Toby was arrested by the policeman')*, there was still evidence for both priming (anticipatory looks to the crook) and active prediction (anticipatory looks to the policeman), but the contribution active prediction was larger. This finding suggests that active prediction may indeed be time-consuming, as there was more active prediction when there was more time available (longer sentence).

A prominent view in the recent literature holds that people use the language production system for prediction of upcoming information during language comprehension (Dell & Chang, 2013; Huettig, 2015; Pickering & Gambi, 2018; Pickering & Garrod, 2007, 2013). In Pickering and Gambi's account (which is based on Pickering & Garrod, 2013), listeners covertly simulate the speaker's utterance and construct derived speaker intention in order to predict subsequent input using their own language production system. Pickering and Gambi suggest that prediction through production is optional and occurs only when listeners have sufficient time and resources available. However, according to the authors, it is also the most effective and most often correct route to prediction. There is both correlational (Hintz, Meyer, & Huettig, 2017; Mani & Huettig, 2012; Rommers, Meyer, & Huettig, 2015) and causal (Martin, Branzi, & Bar, 2018) evidence that production is indeed (at least in some cases) involved in prediction during comprehension.

Pickering and Gambi's (2018) account is also a dual mechanism account in that it assumes a prediction-by-association route in addition to the prediction-by-production route. This route relies on the listener's perceptual experiences and does not involve the production system. Pickering and

Gambi link prediction-by-association to priming effects, in which activation spreads to related concepts or word-forms irrespective of whether that concept or word is likely to occur next in the bottom-up signal. The two routes can work together and one of the two may be engaged more depending on the situation. For example, the prediction-by-production-route may be engaged less when the listener has relatively few resources available.

Two mechanisms (prediction based on higher level information and prediction via association) may not be enough to do justice to the complexity of prediction in language processing (Huettig, 2015; Mani & Huettig, 2013). Huettig (2015) proposes that prediction in comprehension entails at least: production-, association-, combinatorial- and simulation- based mechanisms, in which combinatorial mechanisms involve multiple linguistic constraints and the building up of higher level meaning, and simulation refers to the perceptual simulation of events using mental imagery. In contrast to recent multi-mechanism accounts of prediction, there are also those that assume only one route to (linguistic and non-linguistic) prediction, assigning an important role to event-knowledge, as basis for generating predictions (Altmann & Mirković, 2009; Metusalem et al., 2012).

Modulating Factors

In order to understand which mechanisms are involved in prediction and to find out whether prediction is a prerequisite for language comprehension, researchers have started to explore which factors modulate predictive processing. Some of the factors modulating prediction are bound to the stimuli such as the cloze probability of a sentence frame (the probability of a particular word completing that frame) (DeLong et al., 2005; Ito, Corley, Pickering, Martin, & Nieuwland, 2016), functional associations (Hintz, 2015), predictability of the context (Lau et al., 2013), non-linguistic context (Coco, Keller, & Malcolm, 2016; Hintz et al., 2017) and available time (Chow, Lau, Wang, & Phillips, 2018). Other factors that modulate prediction are bound to the individual, such as working memory capacity, processing speed, verbal fluency, vocabulary size, executive functioning, literacy, age (Huettig & Janse, 2016; Mani & Huettig, 2012; Rommers et al., 2015; Zirnstein, van Hell, & Kroll, 2018) and importantly, (language) experience (Foucart, 2015; Kaan, 2014; Kuperberg & Jaeger, 2016; Peters, Grüter, & Borovsky, 2015; Phillips & Ehrenhofer, 2015). Different mechanisms involved in prediction may not be affected by these factors to the same extent. For instance, Gambi and Pickering (2018) argue that prediction-by-production requires cognitive resources and time, as opposed to prediction-by-association which is less resource intensive but also less accurate. Also, prediction may not be equally robust on all levels; prediction of word form seems to be found less consistently than semantic prediction (e.g., Ito, Martin, & Nieuwland, 2017a; Nieuwland et al., 2017).

In all accounts of predictive language processing, prediction is in one way or another shaped by prior linguistic and/or non-linguistic experience. For example, in Pickering and Gambi's account, shared background knowledge between a speaker and comprehender is used to derive speaker intention in order to predict-by-production, and prediction-by-association depends the strength of learned associations between words (based on how often two words have been encountered together). Mishra et al. (2012) and Mani and Huettig (2014) provided evidence for the notion that prediction is shaped by experience, showing that literacy affects the ability to anticipate during language comprehension. Clearly, linguistic experience is highly likely to differ between native language (L1) and second language (L2) comprehenders, the topic of this dissertation. Therefore, comprehenders may not anticipate as routinely and as effectively in L2 as in L1 comprehension.

LISTENING IN THE NON-NATIVE LANGUAGE: WHAT IS DIFFERENT?

Using context information to predict upcoming information may be particularly useful for L2 comprehenders, as L2 comprehenders may learn from predictions that are not borne out (Dell & Chang, 2013): based on incorrect predictions they may adapt L2 representations and thereby improve their L2 language skills. However, speech comprehension is notoriously difficult in the L2 compared to the L1. L2 comprehension in both the auditory and visual modality tends to be slower, more effortful and error prone than L1 comprehension (Cook, 1997; Cop, Drieghe, & Duyck, 2015; Hahne, 2001; Schmidtke, 2016; Weber & Broersma, 2012).

There are several factors that can account for these disadvantages, and each of these may in turn interfere with predictive processing. Some disadvantages are temporal (e.g. the L2 processing delay) and others are functional (e.g. the findings that L2 processing is more resource consuming or more often incorrect). For instance, lexical access in L2 listening is slower than in L1 listening (Shook, Goldrick, Engstler, & Marian, 2015) and speech perception in noise suffers more in L2 than in L1 (Lecumberri, Cooke, & Cutler, 2010). Here I describe some of the main factors that are thought to individually or conjointly cause temporal and/or functional processing disadvantages in L2 and how these factors may interfere with prediction.

Factors Underlying L2 Disadvantages

Incoming speech activates a set of potential word candidates that compete for recognition based on the extent to which the input matches stored knowledge of the words (Weber & Scharenborg, 2012). In bilingual listening, the set of competing word candidates is larger (almost twice as large) than in monolingual listening (Weber & Broersma, 2012). The main reason for this is that bilingual listeners do not selectively activate words from the target language during speech perception, but also words from the other language. For example, in an eye-tracking study by Marian and Spivey (Marian & Spivey, 2003), Russian-English bilinguals were given instruction to manipulate objects in a display (e.g. pick up the speaker), where the display could contain the target (speaker) and three unrelated objects or both the target and a phonological within (spear) or between language (spichki, 'matches') competitor (and two unrelated objects). The authors found that both within-language and between-language objects whose names were phonologically similar to the target object were fixated more often than unrelated items.

Weber and Cutler replicated this effect (2004) and also found that cross-lingual competition is asymmetric, with competition being larger for L2 listeners. Sentence context modulates these cross-lingual competition effects: When the competing L1 word is incongruent with the sentence context, competitor activation is reduced (Chambers & Cooke, 2009; FitzPatrick & Indefrey, 2010; Lagrou, Hartsuiker, & Duyck, 2013). Lexical competition in L2 word recognition is also thought to be increased because bilinguals have more difficulty deactivating unintended word candidates in L2 (Weber & Broersma, 2012). This is illustrated by Rüschemeyer, Nojack, and Limbach (2008), who found evidence that in L2 processing of words like *roof* (semantically related to *house*) was different when it was preceded by *mouse* (a close phonological neighbor of *house)* than when it was preceded by the unrelated word *lamp.* No such interference was seen in L1. Besides word recognition, L2 speech segmentation (Cutler, Mehler, Norris, & Segui, 1986) and syntactic parsing (Rankin, 2014) are also subject to cross-lingual interference effects.

In addition, the set of competing word candidates during auditory word recognition is larger in bilinguals' L2 than in L1 because of difficulty distinguishing L2 phoneme contrasts that do not exist in participants L1. For example, Dutch listeners often have difficulty distinguishing between the English phonemes /æ/ (as in *hat*) and /ɛ/ (as in *desk)*. This may enlarge the set of word candidates competing for recognition, because for Dutch-English bilinguals, the first syllable of *panda* is compatible not only with *panda* and *panic*, but also with *pencil* and *penny* (Weber & Cutler, 2004). Thus, in spoken L2 perception, there is more uncertainty about candidates for recognition at all levels of processing. But besides increased competition, there are other factors likely to affect language processing as well.

Because bilinguals are less often exposed to L2 than to L1, the accuracy and consistency of linguistic representations may also be weaker in L2 (Kaan, 2014). This idea is consistent with the weaker-links hypothesis of bilingual language processing, which states that the links between phonology and semantics are weaker in bilinguals than in monolinguals because bilinguals have had less experience with linguistic representations in both their languages (as they necessarily divide frequency of use between two languages) (Gollan, Montoya, Cera, & Sandoval, 2008). Also related to accuracy of stored representations, some studies suggest that semantic representations may be less detailed (i.e. contain less features or senses) in L2 than in L1 (Finkbeiner, Forster, Nicol, & Nakamura, 2004; Schoonbaert, Duyck, Brysbaert, & Hartsuiker, 2009). Lower quality representations and weaker links between representations due to lower frequency of use may manifest in processing in several ways. For example, high-frequency words are recognized faster than low-frequency words (e.g., Rubenstein, Garfield, & Millikan, 1970; Scarborough, Cortese, & Scarborough, 1977) and L2 words function as L1 words of lower frequency in recognition because they are practiced less often (e.g., Duyck, Vanderelst, Desmet, & Hartsuiker, 2008). Frequency effects are therefore larger in the non-dominant than in the dominant language. Impoverished semantic representations in L2 affect semantic processing because an L1 word would activate more semantic features than the translation equivalent L1 word (Schoonbaert et al., 2009).

Bilinguals are not only often exposed to L2 representations less frequently, but they also learned and use L2 in a different context (e.g. home versus classroom). Therefore, stored frequency information for linguistic representations and combinations of representations is likely to differ between L1 and L2. Thus, word combinations are not always simply encountered less frequently in the non-dominant L2, but particular words or combinations of words may actually be much more frequent in L2 than in L1. An obvious example of this situation is an idiom that exist in only one of the bilingual's languages (e.g. *To kick the bucket* exists in English and not in Dutch). In addition, the context in which a language is learned (but also the order of acquisition, dominance and age of acquisition) may affect the perceived emotionality in L2 compared to L1 (Pavlenko, 2012).

Finally, the brain may adopt different processing strategies when dealing with L1 or L2 input. For example, bilinguals may rely more on contextual cues in L2 than in L1 (Bradlow & Alexander, 2007; Navarra & Soto-Faraco, 2007). In Navarra and Soto-Faraco for example, Spanish-Catalan bilinguals were able to distinguish the Catalan /e/–/ɛ/phoneme contrast when it was presented audiovisually (lip movements and accompanying speech sound), but not when it was only presented auditorally. Catalan-Spanish bilinguals on the other hand were able to distinguish the contrast in both conditions. Also, bilinguals listening in L2 may have a different syntactic parsing strategy than in L1. Specifically, a prominent account of L2 sentence processing states that L2 listeners have 'shallow syntax'(Clahsen & Felser, 2006). This term is used to indicate that L2 listeners seem to assign different weights to two routes of syntactic processing. Specifically, compared to L1 listeners, L2 listeners are more likely to use a superficial route to derive a syntactic interpretation in L2 than a complex route that makes a full syntactic analysis.

Potential L2 Effects on Prediction.

Each of the factors (increased interference, weaker representations, stored frequency/transitional probability information, and processing strategy) described above may cause delays or increases in required resources in L2, compared to L1 spoken language processing. And each factor may also directly or indirectly interfere with prediction during L2 listening comprehension. For instance, weaker (or even incorrect) lexical representations and increased competition may cause lexical access to be slower (or even to fail). This may hinder the construction of higher level meaning used to predict an upcoming word, or the retrieval of the to be predicted word itself. Weaker links between representations may similarly slow down retrieval of to be predicted words, or a word may activate an associated concept in the L1 but not in the L2. The disadvantages related to L2 processing may each increase the amount of time and resources required for processing, leading to a decrease in the time and resources available for prediction. This could be especially detrimental for resource intensive prediction-by-production, and it may lead to a strategy shift, with predictive processing relying less on prediction-by-production in the L2 than in the L1. In some cases, L2 processing may simply be to slow or resource consuming for prediction to occur at all. In others, predictions may be weaker or restricted to higher levels (such as semantics) in L2.

To add even more complexity to the issue, it may also go the other way around. Prediction in L2 may be enhanced by increased reliance on non-linguistic context information (Bradlow & Alexander, 2007; Navarra & Soto-Faraco, 2007), and increased inhibitory control in bilinguals (Woumans, Ceuleers, Van der Linden, Szmalec, & Duyck, 2015) may attenuate costs associated with prediction errors (Zirnstein et al., 2018). Finally, L2 comprehenders may benefit from predictive processing if they learn from incorrect predictions by adjusting representations and the links between them after encountering unexpected input (Dell & Chang, 2013). Thus, there is a complex interplay of factors that could potentially modulate predictive processing in the L2. Kaan (2014) suggests that the mechanisms underlying prediction are essentially the same in L2 as in L1, but that individual differences (that also affect prediction in L1) may impact prediction in L2 differently.

Models of Bilingual Language Processing

Most models of bilingual language processing do not incorporate predictive pre-activation. Also, most models are either not designed for language processing in a particular modality (e.g. the Revised Hierarchical Model; Kroll & Stewart, 1994; the Distributed Feature Model; Van Hell & De Groot, 1998), or they are focused exclusively on the visual modality (Bilingual Interactive Activation Model Plus; Dijkstra & van Heuven, 2002; Multilink; Dijkstra et al., 2018). The only model that does particularly focus on the auditory modality is The Bilingual Language Interaction Network for Comprehension of Speech model (BLINCS) (Shook & Marian, 2013). BLINCS is a connectionist model consisting of an interconnected network of dynamic, self-organising maps. The model assumes four levels of representation: phonological, phono-lexical, ortho-lexical and semantic. All connections between levels are bidirectional. The model assumes an integrated L1/L2 lexicon in which the two languages are separated into regions according to phono-tactic probabilities of the input. Conceptual representations are shared across languages in BLINCS, although the authors note that conceptual representations across languages may not always be exactly the same. Also, even if conceptual representations are the same, the strength of connections between concepts may potentially differ between languages. An interesting feature of the model is that it accounts not only for effects of the auditory bottom-up input, but also takes into account effects of visual (non-linguistic) input, such as from a scene in the visual world paradigm. This information directly feeds into the semantic level and can thereby constrain lexical activation. Unfortunately, as the authors note, also this model has yet to be extended to incorporate prior activation from the linguistic context and effects of expectations based on context information. Therefore, it does not yet lead to predictions about prediction during L2 speech comprehension, nor about the mechanisms underlying it.

PREDICTION IN NON-NATIVE LANGUAGE COMPREHENSION

There is an increasing body of evidence showing that predictive processing occurs in the L2 in the visual (Foucart, Martin, Moreno, & Costa, 2014) and auditory (Chambers & Cooke, 2009;

Foucart, Ruiz-Tada, & Costa, 2015; Hopp, 2013, 2015; Ito, Corley, & Pickering, 2017; Ito, Pickering, & Corley, 2018) modality. However, there is also evidence suggesting that prediction is sometimes weaker or even absent in L2 comprehension (Hopp, 2015; Ito et al., 2018; Kaan, Kirkham, & Wijnen, 2014; Martin et al., 2013; Sagarra & Casillas, 2018; van Bergen & Flecken, 2017). Whether or not prediction effects are found in L2 may depend on factors such as the level of processing, L2 listener proficiency (e.g., Dussias, Valdés Kroff, Guzzardo Tamargo, & Gerfen, 2013; Hopp, 2013), and L1-L2 language similarity (van Bergen & Flecken, 2017).

L1-L2 Similarity and L2 Proficiency

L1-L2 relatedness may be a factor determining whether or not prediction effects are found in L2 comprehension. Van Bergen and Flecken manipulated cross-linguistic similarity more directly, by comparing groups of bilinguals with different L1's. The groups differed with respect to familiarity with placement verbs specifying object position. German, like Dutch specifies object position in placement verbs; a different verb is used for *put* when the relevant object's end position is lying down (*leggen,* 'put.LIE') than when it is placed standing up (*zetten*, 'put.STAND'), whereas English (*put*) and French do not (*mettre*). Participants were exposed to sentences such as *de jongen zette/legde/plaatste kort geleden een bal/taart/fles op de tafel 'the boy put.stand/put.lie/put recently a ball/cake/bottle on the table'* while they looked at displays containing an object in lying position (e.g. ball), an object in standing position (e.g. cake) and one object depicted both in standing and in lying position (bottle). Indeed, German-Dutch bilinguals, like Dutch native speakers, launched anticipatory eye-movements to the objects corresponding to the position encoded by the verb, whereas French-Dutch and English-Dutch bilinguals did not. The authors interpret the findings in terms of linguistic experience, and argue that the amount of linguistic experience determines the automaticity of (predictive) processing.

L2 proficiency also influences predictive processing in L2. For instance, in Hopp (2013), only participants with native-like mastery of L2 gender assignment were able to use article gender information as cue for prediction. Also, Sagarra and Casillas (2018) recently showed that advanced learners but not beginning learners of Spanish employed prosodic information to anticipate word suffixes. Peter, Grüter and Borovsky show that proficiency may not only impact whether or not prediction occurs but that it may also affect prediction strategy. In their visual world study, low-proficient and high-proficient non-native comprehenders listened to sentences (e.g., *"The pirate chases the ship"*) while they looked at displays featuring agent-related, action-related and unrelated pictures. High-proficient bilinguals were faster than low-proficient bilinguals. In addition, the low-proficient bilinguals were more likely than high-proficient bilinguals to anticipate locally-coherent action-related distractors (e.g., a cat). The authors suggest that low-proficient bilinguals adapted to a higher level of uncertainty in interpretation by activating less likely but locally coherent candidates. Another interpretation would be that the low-proficient bilinguals relied more on untargeted prediction-by-association and less on prediction-by-production.

Word Form and Syntax Levels

Whether or not L1-L2 differences in predictive processing our found may also depend on the level of processing. To our knowledge, there is no compelling evidence to date suggesting that bilinguals predict information on the word form level. Ito et al. (2018) studied this behaviourally. Native English and Japanese-English bilinguals listened to constraining sentences in English (*e.g. The tourists expected rain when the sun went behind the ...*), and looked at displays containing either a target object (*cloud*; in Japanese: *Kumo*), a phonological competitor for the target object name in English (*clown*), a phonological competitor for the target object name in Japanese (bear; *kuma*), or an unrelated object (*globe*; *tikyuugi*). Native listeners fixated target objects and English competitors more than distractor objects. Non-native listeners only fixated targets more often than distractors (though later than the native listeners), and not English or Japanese phonological competitors, indicating that they predicted target word semantics but not word form. There is also no neural evidence for pre-activation of word form: EEG studies focusing on sentence reading have failed to find evidence for prediction of word form in bilinguals in L2 (Ito, Martin, & Nieuwland, 2017b; Martin et al., 2013; but note that recently, Nieuwland et al., 2017, in a multi-lab study, also failed to replicate prediction of word-form in native-speakers). In an EEG study in the visual domain by Martin et al. (2013) for instance, native speakers of English and late Spanish–English bilinguals read sentences in English with predictable or less predictable sentence-final nouns. Event-related potentials were measured at the article preceding the sentence final noun. The article was always congruent with the final noun, but not always with the expected noun (e.g., *Since it is raining, it is better to go out with an umbrella* $_{[EXPECTED]}$*/ a raincoat* $_{[UNEXPECTED]}$. If participants indeed predicted umbrella, a semantic anomaly effect should be elicited by the article *a* relative to *an*, because a is incongruent with umbrella. Thus, the target for prediction is the lexical form and the congruent article. Martin et al. indeed found an N400-effect for the unexpected versus the expected nouns in

L1 and L2 readers. The N400 effect was also significant for the article in L1 readers, but not in L2 readers. Thus, in this study there was no evidence for prediction of word form in L2 readers either.

A number of studies have found weaker or no prediction effects when syntactic information in involved (either as cue for prediction or as predictee) (Hopp, 2013, 2015; Kaan et al., 2014; Mitsugi & Macwhinney, 2015). Hopp (2015) specifically contrasted prediction based on (morpho-)syntactic cues and prediction based on lexical-semantic cues in a visual world paradigm. Native German participants and English-German bilingual participants looked at a scene depicting three possible actors and one control object while they listened to SVO (e.g. *The$_{[nom]}$ wolf kills soon the$_{[acc]}$ deer*) or OVS (e.g., *The$_{[acc]}$ wolf kills soon the$_{[nom]}$ hunter*) sentences in German. Anticipatory looks were found to expected patients (*the deer)* before the onset of the second NP in SVO sentences and at expected agents (*the hunter*) in OVS sentences in the native listener group. In contrast, the English-German bilinguals were more likely to look at patient objects before the onset of the second NP, both when the first NP had nominative or accusative case marking. The findings show that whereas L2 listener seemed to anticipate the second NP based on the meaning of the first NP and the verb, they were unable to employ case-marking information to adjust their prediction. Similarly, Mitsugi and Macwhinney showed that native speakers of Japanese used case marking information to anticipate an upcoming constituent whereas learners of Japanese did not. In contrast, in another visual world study, Hopp (2013) showed that English-German bilinguals anticipated target objects whose syntactic gender agreed with a spoken article, but only in bilinguals with native-like mastery of German gender assignment in production. Neural evidence also suggests that L2 listeners can predict nouns and their syntactic gender in reading (Foucart et al., 2014) and listening (Foucart et al., 2015), at least when the bilinguals' languages have similar gender-noun agreement rules.

Semantic Level

In contrast to word-level and syntactic prediction, semantic prediction is often assumed to be intact in bilinguals, and some of the behavioral and EEG evidence indeed suggests that it is. As discussed above, Hopp (2015) showed that L2 listeners do not employ case-marking information in predictive processing (like L1 listeners do), but like L1 listeners, they employ lexical-semantic information in predictive processing. In another visual world study, Ito et al. (2017) used a paradigm similar to (Altmann & Kamide, 1999) in which the verbs were manipulated to restrict the subsequent possible referents in the display (e.g. *The lady will fold/find the scarf*, with the scarf being the only foldable object in a four-picture display). Both L1 English listeners and bilinguals with English as L2 (various L1's), used semantic information provided by the verb to anticipate upcoming referents. In addition, anticipatory eye-movements were equally affected by an additional cognitive load (remembering words) in L1 and L2, showing that cognitive resources are implicated in predictive processing in the visual world paradigm. Hintz and Meyer (2015) used a novel version of the visual world paradigm, in which Dutch L1 comprehenders and German-Dutch bilinguals were shown a display featuring a clock while they listened to simple mathematical equations in Dutch (e.g., *three plus eight is eleven*). In the comprehension condition participants listened to the entire equation including the solution, whereas in the other condition participants only heard the equations upto the solution and had to provide the solutions themselves. Both L1 listeners and L2 listeners fixated the solution well before hearing it in the comprehension condition, and before producing it in the production condition. L2 listeners were only slightly slower than L1 listeners. There is also evidence from EEG studies showing that bilinguals anticipate lexical-semantic information based on lexical-semantic information from the sentence context in L2 reading and listening. Foucart et al. (2014) used a paradigm similar to Martin et al. (2013) but manipulated gender congruency of an article with an expected noun to elicit an N400 effect (e.g., *The pirate had the secret map, but he never found the$_{[masc]}$ treasure$_{[EXPECTED]}$/ the$_{[fem]}$ cave $_{[UNEXPECTED]}$ he was looking for).* An N400 effect was elicited by the article incongruent with the expected word (though always congruent with the sentence final word) compared to the article congruent with the expected word in Spanish monolinguals, French-Spanish bilinguals and, Spanish-Catalan bilinguals. This finding suggests that bilinguals are able to predict upcoming words based on lexical-semantic information from the sentence context. In a follow-up study using a similar paradigm, Foucart et al. (2015) replicated this finding in listening in French-Spanish late bilinguals. All critical nouns were muted in this study. Interestingly, in a subsequent recognition test, expected words were falsely recognized as having been heard more often than unexpected words, suggesting that a memory trace of expected words was created.

Even though semantic prediction effects have been found in L2 like in L1, there is reason to expect that in more challenging conditions, prediction effects will differ between L1 and L2. For one, representations of L2 words may semantically poorer than representations of L1 words (Finkbeiner et al., 2004; Schoonbaert et al., 2009). Therefore, two words that share semantic features in L1 may have no or fewer shared features in L2. If so, semantic predictions based on lexical-semantic information from the sentence context should be affected by language, just like predictions based on morpho-syntactic information. Perhaps related to this hypothesis, weaker links between word forms and semantics due to less practice in L2 than in L1 (Gollan et al., 2008) could also hinder information retrieval in lexical-semantic prediction. There are indeed some indications that lexical-semantic prediction is not always intact in L2. For instance, Japanese-English bilinguals listening to constraining sentences showed anticipatory eye-movements to a predictable target object later than English native speakers (e.g., *cloud*, when listening to *The tourists expected rain when the sun went behind the* . . .) (Ito et al., 2018). Also, using EEG, Ito et al. (2017b) found an attenuation of the N400 elicited by a semantic competitor (*page*) of a predictable target word (*book*) (following *The student is going to the library to borrow a*…). However, the attenuation did not depend on cloze probability, and therefore the authors did not interpret the effect as evidence for semantic pre-activation. Using the same paradigm, semantic pre-activation was found in native readers (Ito, Corley, Pickering, Martin, & Nieuwland, 2016).

Taken together, the evidence suggests that bilinguals can predict during comprehension, but that they often do not do so to the same extent as native comprehenders. Proficiency and L1-L2 similarity seem to play a role in whether or not prediction effects are found to be intact or not in L2. The level of processing may also be a factor determining the probability of successful prediction. Whereas there is no evidence that bilinguals predict word form, and bilinguals do not consistently predict when syntactic information is involved, semantic prediction is often (though not always) found to be intact. This is in line with Pickering and Gambi's (2018) hypothesis that prediction-by-production proceeds in the same order as actual production (first semantics, then syntax and finally form). If prediction-by-production is weaker or delayed in L2 than in L1, this should be most pronounced in prediction involving word form and syntactic information. Pickering and Gambi suggest that comprehenders may rely less on prediction-by-production in L2 than in L1. If so, prediction-by-production is expected to be weaker in L2 than in L1, whereas prediction via low-level lexical associations is expected to be largely intact in L2. So far, many studies have looked for whether or not prediction effects could be found in L2, and they did not directly compare prediction effects in L1 and L2. In addition, a lot of research has focussed on syntactic and word-form prediction. Prediction at those levels may be more language dependent than semantic prediction, as semantic representations are often assumed to be (mostly) language independent (e.g., Dijkstra & van Heuven, 2002; Kroll & Stewart, 1994; Shook & Marian, 2013). Therefore, effects of (non-)nativeness on syntactic and word-form prediction may depend more strongly on L1/L2 cross-linguistic similarity, compared to lexical-semantic prediction. Therefore, in this dissertation we focussed on comparing semantic prediction in L1 and L2.

SPEECH PRODUCTION AFTER SPEECH PERCEPTION IN THE NON-NATIVE LANGUAGE

A number of prominent accounts of prediction in language comprehension assume that prediction involves the speech production system (Dell & Chang, 2013; Pickering & Gambi, 2018; Pickering & Garrod, 2013), and that there is parity between representations in comprehension and production. Such accounts entail that perception of an interlocutor's speech during comprehension can cause subsequent adaptations in speech perception (e.g., Norris, McQueen, & Cutler, 2003), but also in speech production. Therefore, here we also assessed interactions between speech comprehension and speech production.

According to Pickering and Garrod's alignment model (Garrod & Pickering, 2004, 2009), adaptations in speech production after speech perception serve the purpose of optimizing mutual understanding between interlocutors. Gambi and Pickering (2013) have suggested that (phonetic) adaptation occurs because listeners covertly imitate the speaker using their own language production system, in order to generate predictions about upcoming speech. There is indeed a large body of evidence suggesting that adaptation occurs at the syntactic (e.g., Bernolet, Hartsuiker, & Pickering, 2012, 2013; Pickering & Branigan, 1999), lexical (e.g., Branigan, Pickering, Pearson, McLean, & Brown, 2011), and phonetic (e.g., Babel, 2012; Lametti, Krol, Shiller, & Ostry, 2014; Pardo, 2006) levels.

Phonetic adaptation may also serve as a useful L2 learning strategy when a bilingual interacts in her L2 with a speaker that is more proficient (Costa, Pickering, & Sorace, 2008). On the other hand, weaker representations lack of automaticity in L2 production may hamper efficient adaptation. This may be particularly so when speakers are very different from each other, such as a native and a non-native speaker. Gambi and Pickering suggest, that in such cases, prediction-by-simulation (using the production system) may fail because the listener does not have enough experience to imitate the native speaker. In the case where interlocutors perceive themselves as being very different from each other, they may rely more on the other route to prediction during comprehension: prediction-by-association. This route does not involve the production system and therefore adaptations in production is not expected.

There is some evidence suggesting that non-native speakers adapt their speech production when interacting with native or non-native speakers (Hwang, Brennan, & Huffman, 2015; Kim, Horton, & Bradlow, 2011; Trofimovich & Kennedy, 2014), but most studies have used subjective similarity ratings instead of acoustical measures for specific target sounds. Also, it remains unclear whether adaptation effects for specific phonemes last during (or even after) conversation or wither the effects decay quickly. Finally, if adaptation indeed serves the purpose of aligning situation models and thereby enhancing conversation, then the extent to which speakers engage in adaptation may depend on social context, for instance, whether the other speaker is physically present or not. Babel (2012) showed that participants adapted more to speech over headphones in a shadowing task (in L1) when there was a picture of the speaker presented on the screen in front of the participant then when there was no picture on the screen. As comprehenders tend to rely more on context information in L2 than in L1 (Bradlow & Alexander, 2007; Navarra & Soto-Faraco, 2007), social context may be particularly relevant for an L2 speaker interacting with another speaker.

CURRENT DISSERTATION

An increasing number of studies have investigated whether people predict upcoming information when listening to speech in L2 like they do in L1. So far, the results have been inconsistent. Some studies have found evidence for weaker, slower, or no prediction at all in L2 (Hopp, 2015; Ito et al., 2018; Martin et al., 2013; Mitsugi & Macwhinney, 2015; Sagarra & Casillas, 2018; van Bergen & Flecken, 2017), whereas other find prediction in L2 like in L1 (Foucart et al., 2015; Hopp, 2015; Ito, Corley, et al., 2017). If Pickering & Gambi's hypothesis that comprehenders rely less on prediction-by-production is correct, differences between prediction in L1 and L2 are expected to arise mainly when prediction involves syntactic or word form information, whereas predictions based on low-level lexical associations should be largely intact. Previous research has shown differences between L1 and L2 predictive processing of word form and syntax. In the current dissertation we focussed on lexical-semantic prediction in bilinguals. Subtle differences in semantic prediction in L1 and L2 are expected because of differences in the structure of L1 and L2 semantic memory (e.g., poorer representations and weaker links between phonology and semantics) due to differences in linguistic experience. Importantly, we also assessed mechanisms that potentially underlie L1-L2 differences in semantic prediction, when it is found. Specifically, we studied the role of availability of cognitive resources and of processing speed in L1 and L2 predictive processing.

In **CHAPTER 2-4** we studied prediction of semantics based on the lexical-semantic sentence context in L1 and L2 using the visual world paradigm. The visual world paradigm was first employed by Cooper (1974), and it began to be used on a larger scale after publication of a study by Tanenhaus et al. (1995). The paradigm has the advantage of tracking language activation in real time; research has shown that object fixations in the visual world are closely time-locked to lexical access (Allopenna, Magnuson, & Tanenhaus, 1998). In addition, there is no need for an extra-linguistic task that may confound the results. Prediction was compared across languages within the same individuals. This way, we eliminated confounding effects of life experience and individual cognitive differences that may affect prediction such as working memory, processing speed (Huettig & Janse, 2016), age (Federmeier & Kutas, 2005), and verbal fluency (Rommers, Meyer, & Huettig, 2015). This also eliminates the high inter-individual variability that characterizes eye movements (Bargary et al., 2017; Rayner, 1998) and which may confound between-group differences in visual world paradigms. This method entails that language processing is compared across two different languages. To deal with language differences, we either matched L1 and L2 stimuli on a number of properties (such as frequency and length), or those properties were included as factors in the analyses.

CHAPTER 2 focused on investigating whether bilinguals anticipated upcoming referents based on information extracted at the verb in their L1 and L2, in simple subject-verb-object (SVO) sentences. The paradigm was based on the seminal visual world study by Altmann and Kamide (1999), in which the verb restricted the subsequent domain of reference. The first aim of this study was to test whether prediction occurred at all in L2 listening, and the second aim was to compare semantic prediction in L2 directly to semantic prediction in L1. In addition, a monolingual control group was tested to see whether any differences between prediction in L1 and L2 were due to the language manipulation (English vs. Dutch) or due to language status (L1 vs. L2). Stimulus characteristics were carefully matched between languages. Although in this paradigm predictions could theoretically be generated based on higher level information, it is also likely that association-based mechanisms were involved to a large extent, as predictions were based on semantic information from only the verb. If Pickering and Gambi's (2018) hypothesis that prediction-by-association is mostly intact in language learners is correct, no major difference between prediction in L1 and L2 is expected here.

Then, in **CHAPTER 3**, we studied semantic prediction in bilinguals in a more fine-grained way. We hypothesized that subtle differences between semantic prediction in L1 and L2 should arise in more challenging conditions. That is, when target predictions were more likely to be based on higher order (message level) information. This type of prediction is likely to require more resources (possibly unavailable in L2). Here, we used a more naturalistic and larger set of sentences to test whether differences between prediction in L1 and L2 occurred in more demanding situations. The sentences were longer and syntactically more complex. The picture preview time was very short. Further, we also tested whether spreading of semantic activation differed between L1 and L2 by adding a condition with a semantic competitor of the target for prediction. The semantic distance between target-competitor pairs was included as factor in the analyses. Based on Pickering and Gambi (2018)'s hypothesis that prediction-by-association is automatic and not optional, pre-activation of semantic competitors was expected to be relatively intact in L2. However, weaker links in L2 than in L1 due to lower frequency of use could result in a language difference in the impact of semantic distance on competitor pre-activation.

Differences between semantic prediction in L1 and L2 are not found consistently, and it remains unclear what mechanisms underlie the difference, when it is found. **CHAPTER 4** focussed on the factors that potentially underlie L1-L2 difference in semantic prediction. One potential factor is the cognitive load associated with L2 processing. L2 processing seems to require more cognitive resources than L1 processing (Abutalebi, 2008; Francis & Gutiérrez, 2012; McDonald, 2006). Pickering and Gambi (2018) have suggested that prediction-by-production (particularly the later stages) may be impaired in populations with limited availability of resources, as this route to production requires resources and time. We therefore hypothesized that if prediction in L2 is weaker because of the higher cognitive load associated with L2 processing, an additional cognitive load would be particularly detrimental for prediction in L2. The second potential factor, processing speed, was chosen using a similar line of reasoning. L2 processing is slower compared to L1 processing (Cop, Keuleers, Drieghe, & Duyck, 2015; Hahne, 2001; Moreno, Rodríguez-Fornells, & Laine, 2008). As prediction-by-production requires time, slower processing may underlie the difference between prediction in the L1 and the L2. This hypothesis entails that slowing down speech input in L2 should enhance prediction, and that speeding up L1 input should attenuate prediction. In two experiments we manipulated cognitive load and stimulus presentation rate to test these hypotheses. This study used the same materials as **CHAPTER 3** and it was therefore also an attempt to replicate the findings of our previous experiment.

Recent accounts of prediction assume representational parity between production and comprehension (Pickering & Gambi, 2018; Pickering & Garrod, 2013). If this assumption is correct, adaptations in perception based on incorrect predictions can affect subsequent speech production as well. The extent to which comprehenders rely on prediction-by-production (and thus the extent to which the production system is involved during comprehension) may depend on social variables such as perceived similarity with the speaker. This entails that, prediction and subsequent adaptation in production may fail to occur if interlocutor pairs are highly dissimilar (Gambi & Pickering, 2013). Therefore, in **CHAPTER 5** we tested whether listening to speech in the L2 produced by a native speaker, leads to changes in subsequent L2 production. In this study we asked participants to read aloud sentences in L2 containing two target phonemes, before and after exposure to a native speaker producing sentences with the same phonemes. There was a confederate present condition (native speaker was present in the room with the participant) and a confederate absent condition to see whether amount of adaptation depended on social context. We ran acoustic analyses of the recordings of participant and confederate utterances.

REFERENCES

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica*, *128*(3), 466–478. https://doi.org/10.1016/j.actpsy.2008.03.014

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*(4), 419–439. https://doi.org/10.1006/jmla.1997.2558

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*(4), 502–518. https://doi.org/10.1016/j.jml.2006.12.004

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and Prediction in Human Sentence Processing. *Cognitive Science*, *33*(4), 583–609. https://doi.org/10.1111/j.1551-6709.2009.01022.x

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, *40*(1), 177–189. https://doi.org/10.1016/j.wocn.2011.09.001

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*(7), 280–289. https://doi.org/10.1016/j.tics.2007.05.005

Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1235–1243. https://doi.org/10.1098/rstb.2008.0310

Bargary, G., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). Individual differences in human eye movements: An oculomotor signature? *Vision Research*, *141*, 157–169. https://doi.org/10.1016/j.visres.2017.03.001

Boland, J. E. (2005). Visual arguments. *Cognition*, *95*(3), 237–274.

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*, *34*(6), 774–806. https://doi.org/10.1006/jmla.1995.1034

Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America*, *121*(4), 2339. https://doi.org/10.1121/1.2642103

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 1029–1040. https://doi.org/10.1037/a0015901

Chow, W.-Y., Lau, E., Wang, S., & Phillips, C. (2018). Wait a second! delayed impact of argument roles on on-line verb prediction. *Language, Cognition and Neuroscience*, *0*(0), 1–26. https://doi.org/10.1080/23273798.2018.1427878

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(1), 3–42. https://doi.org/10.1017/S0142716406060024

Coco, M. I., Keller, F., & Malcolm, G. L. (2016). Anticipation in Real-World Scenes: The Role of Visual Context and Visual Memory. *Cognitive Science*, *40*(8), 1995–2024. https://doi.org/10.1111/cogs.12313

Cook, V. (1997). The consequences of bilingualism for cognitive processing. In A. Groot & J. F. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 279–300). Lawrence Erlbaum.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*(1), 84–107. https://doi.org/10.1016/0010-0285(74)90005-X

Cop, U., Drieghe, D., & Duyck, W. (2015). Eye movement patterns in natural reading: A comparison of monolingual and bilingual reading of a novel. *PloS One*, *10*(8), e0134008.

Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015). Frequency effects in monolingual and bilingual natural reading. *Psychonomic Bulletin & Review*, *22*(5), 1216–1234. https://doi.org/oi: 10.3758/s13423-015-0819-2

Costa, A., Pickering, M. J., & Sorace, A. (2008). Alignment in second language dialogue. *Language and Cognitive Processes*, *23*(4), 528–556. https://doi.org/10.1080/01690960801920545

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, *25*(4), 385–400. https://doi.org/10.1016/0749-596X(86)90033-1

Dell, G. S., & Chang, F. (2013). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120394–20120394. https://doi.org/10.1098/rstb.2012.0394

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121. https://doi.org/10.1038/nn1504

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal of Memory and Language*, (41), 496–518.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(03). https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2018). Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, 1–23. https://doi.org/10.1017/S1366728918000287

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). WHEN GENDER AND LOOKING GO HAND IN HAND. *Studies in Second Language Acquisition*, *35*(02), 353–387. https://doi.org/10.1017/S0272263112000915

Duyck, W., Assche, E. V., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 663–679. https://doi.org/10.1037/0278-7393.33.4.663

Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R. J. (2008). The frequency effect in second-language visual word recognition. *Psychonomic Bulletin & Review*, *15*(4), 850–855. https://doi.org/10.3758/PBR.15.4.850

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, *24*(6), 409–436. https://doi.org/10.1007/BF02143160

Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, *42*(2), 133–141. https://doi.org/10.1111/j.1469-8986.2005.00274.x

Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language*, *51*(1), 1–22. https://doi.org/10.1016/j.jml.2004.01.004

FitzPatrick, I., & Indefrey, P. (2010). Lexical competition in nonnative speech comprehension. *Journal of Cognitive Neuroscience*, *22*(6), 1165–1178. https://doi.org/10.1162/jocn.2009.21301

Foreign language skills statistics - Statistics Explained. (2015). Retrieved May 23, 2018, from http://ec.europa.eu/eurostat/statistics-explained/index.php?title=Foreign_language_skills_statistics

Foucart, A. (2015). Prediction is a question of experience. *Linguistic Approaches to Bilingualism*, *5*(4), 465–469. https://doi.org/10.1075/lab.5.4.04fou

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1461–1469. https://doi.org/10.1037/a0036756

Foucart, A., Ruiz-Tada, E., & Costa, A. (2015). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition*, 1–7. https://doi.org/10.1017/S1366728915000486

Francis, W. S., & Gutiérrez, M. (2012). Bilingual recognition memory: Stronger performance but weaker levels-of-processing effects in the less fluent language. *Memory & Cognition*, *40*(3), 496–503. https://doi.org/10.3758/s13421-011-0163-3

Gambi, C., & Pickering, M. J. (2013). Prediction and imitation in speech. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00340

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11. https://doi.org/10.1016/j.tics.2003.10.016

Garrod, S., & Pickering, M. J. (2009). Joint Action, Interactive Alignment, and Dialog. *Topics in Cognitive Science*, *1*(2), 292–304. https://doi.org/10.1111/j.1756-8765.2009.01020.x

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Grosjean, F. (2010). *Bilingual: life and reality*. Cambridge, Mass: Harvard University Press.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, *30*(3), 251–266. https://doi.org/10.1023/A:1010490917575

Hintz, F. (2015). *Predicting language in different contexts: the nature and limits of mechanisms in anticipatory language processing*. Radboud Universiteit Nijmegen, Nijmegen. Retrieved from http://hdl.handle.net/2066/147122

Hintz, F., & Meyer, A. S. (2015). Prediction and Production of Simple Mathematical Equations: Evidence from Visual World Eye-Tracking. *PLOS ONE*, *10*(7), e0130766. https://doi.org/10.1371/journal.pone.0130766

Hintz, F., Meyer, A. S., & Huettig, F. (2017). Predictors of verb-mediated anticipatory eye movements in the visual world. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(9), 1352–1374. https://doi.org/10.1037/xlm0000388

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*(1), 33–56. https://doi.org/10.1177/0267658312461803

Hopp, H. (2015). Semantics and Morphosyntax in Predictive L2 Sentence Processing. *International Review of Applied Linguistics.* https://doi.org/10.1515/iral-2015-0014

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459

Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, *31*(1), 19–31. https://doi.org/10.1080/23273798.2015.1072223

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*(2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003

Hwang, J., Brennan, S. E., & Huffman, M. K. (2015). Phonetic adaptation in non-native spoken dialogue: Effects of priming and audience design. *Journal of Memory and Language*, *81*, 72–90. https://doi.org/10.1016/j.jml.2015.01.001

Ito, A., Corley, M., & Pickering, M. J. (2017). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, *21*(2), 251–264. https://doi.org/10.1017/S1366728917000050

Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language*, *86*, 157–171. https://doi.org/10.1016/j.jml.2015.10.007

Ito, A., Martin, A. E., & Nieuwland, M. S. (2017a). How robust are prediction effects in language comprehension? Failure to replicate article-elicited N400 effects. *Language, Cognition and Neuroscience*, *32*(8), 954–965. https://doi.org/10.1080/23273798.2016.1242761

Ito, A., Martin, A. E., & Nieuwland, M. S. (2017b). On predicting form and meaning in a second language. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*, *43*(4), 635–652. https://doi.org/10.1037/xlm0000315

Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, *98*, 1–11. https://doi.org/10.1016/j.jml.2017.09.002

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, *4*(2), 257–282. https://doi.org/10.1075/lab.4.2.05kaa

Kaan, E., Kirkham, J., & Wijnen, F. (2014). Prediction and integration in native and second-language processing of elliptical structures. *Bilingualism: Language and Cognition*, 1–18. https://doi.org/10.1017/S1366728914000844

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133–156. https://doi.org/10.1016/S0749-596X(03)00023-8

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, *2*(1). https://doi.org/10.1515/labphon.2011.004

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. *Cognition*, *95*(1), 95–127. https://doi.org/10.1016/j.cognition.2004.03.002

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*(2), 149.

Kukona, A., Fang, S.-Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The Time Course of Anticipatory Constraint Integration. *Cognition*, *119*(1), 23–42. https://doi.org/10.1016/j.cognition.2010.12.002

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Research*, *1146*, 23–49. https://doi.org/10.1016/j.brainres.2006.12.063

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 190–207).

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161–163. https://doi.org/10.1038/307161a0

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2013). The influence of sentence context and accented speech on lexical access in second-language auditory word recognition. *Bilingualism: Language and Cognition*, *16*(03), 508–517. https://doi.org/10.1017/S1366728912000508

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience*, *25*(3), 484–502. https://doi.org/10.1162/jocn_a_00328

Lecumberri, M. L. G., Cooke, M., & Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication*, *52*(11–12), 864–886. https://doi.org/10.1016/j.specom.2010.08.014

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00226

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 843–847. https://doi.org/10.1037/a0029284

Mani, N., & Huettig, F. (2013). Towards a complete multiple-mechanism account of predictive language processing. *Behavioral and Brain Sciences*, *36*(04), 365–366. https://doi.org/10.1017/S0140525X12002646

Mani, N., & Huettig, F. (2014). Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. *Journal of Experimental Child Psychology*, *126*, 264–279. https://doi.org/10.1016/j.jecp.2014.05.004

Marian, V., & Spivey, M. (2003). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, *24*, 173–193. https://doi.org/10.1017.S0142716403000092

Martin, C. D., Branzi, F. M., & Bar, M. (2018). Prediction is Production: The missing link between language production and comprehension. *Scientific Reports*, *8*(1), 1079. https://doi.org/10.1038/s41598-018-19499-4

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*(3), 381–401. https://doi.org/10.1016/j.jml.2006.06.006

Metusalem, R., Kutas, M., Urbach, T. P., Hare, M., McRae, K., & Elman, J. L. (2012). Generalized event knowledge activation during online sentence comprehension. *Journal of Memory and Language*, *66*(4), 545–567. https://doi.org/10.1016/j.jml.2012.01.001

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye-movements are modulated by reading ability - Evidence from Indian low and high literates. *Journal of Eye Movement Research*, *5*(1), 1–10. https://doi.org/10.16910/jemr.5.1.3

Mitsugi, S., & Macwhinney, B. (2015). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, *19*(01), 19–35. https://doi.org/10.1017/S1366728914000881

Moreno, E. M., Rodríguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics*, *21*(6), 477–508. https://doi.org/10.1016/j.jneuroling.2008.01.003

Navarra, J., & Soto-Faraco, S. (2007). Hearing lips in a second language: visual articulatory information enables the perception of second language sounds. *Psychological Research*, *71*(1), 4–12. https://doi.org/10.1007/s00426-005-0031-5

Nieuwland, M., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., … Huettig, F. (2017). Limits on prediction in language comprehension: A multi-lab failure to replicate evidence for probabilistic pre-activation of phonology. *BioRxiv*, 111807. https://doi.org/10.1101/111807

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*(2), 204–238. https://doi.org/10.1016/S0010-0285(03)00006-9

Otten, M., & Van Berkum, J. J. A. (2008). Discourse-Based Word Anticipation During Language Processing: Prediction or Priming? *Discourse Processes*, *45*(6), 464–496. https://doi.org/10.1080/01638530802356463

Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology*, *47*(6), 405–428. https://doi.org/10.1080/00207594.2012.743665

Peters, R., Grüter, T., & Borovsky, A. (2015). Anticipatory and Locally Coherent Lexical Activation Varies as a Function of Language Proficiency. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 1865–1870). Austin, TX.

Phillips, C., & Ehrenhofer, L. (2015). The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism*, *5*(4), 409–453. https://doi.org/10.1075/lab.5.4.01phi

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: a theory and review. *Psychological Bulletin*. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105–110. https://doi.org/10.1016/j.tics.2006.12.002

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(04), 329–347. https://doi.org/10.1017/S0140525X12001495

Rankin, T. (2014). Variational learning in L2: The transfer of L1 syntax and parsing strategies in the interpretation of *wh* -questions by L1 German learners of L2 English. *Linguistic Approaches to Bilingualism*, *4*(4), 432–461. https://doi.org/10.1075/lab.4.4.02ran

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. http://dx.doi.org/10.1037/0033-2909.124.3.372

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, & Psychophysics*, *77*(3), 720–730. https://doi.org/10.3758/s13414-015-0873-x

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 487–494. https://doi.org/10.1016/S0022-5371(70)80091-3

Sagarra, N., & Casillas, J. V. (2018). Suprasegmental information cues morphological anticipation during L1/L2 lexical access. *Journal of Second Language Studies*, *1*(1), 31–59. https://doi.org/10.1075/jsls.17026.sag

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(1), 1–17. https://doi.org/10.1037/0096-1523.3.1.1

Schmidtke, J. (2016). The Bilingual Disadvantage in Speech Understanding in Noise Is Likely a Frequency Effect Related to Reduced Language Exposure. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00678

Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2009). Semantic and translation priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition*, *37*(5), 569–586. https://doi.org/10.3758/MC.37.5.569

Sedivy, J. C., K Tanenhaus, M., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*(2), 109–147. https://doi.org/10.1016/S0010-0277(99)00025-6

Shook, A., Goldrick, M., Engstler, C., & Marian, V. (2015). Bilinguals Show Weaker Lexical Access During Spoken Sentence Comprehension. *Journal of Psycholinguistic Research*, *44*(6), 789–802. https://doi.org/10.1007/s10936-014-9322-6

Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for Comprehension of Speech. *Bilingualism: Language and Cognition*, *16*(02), 304–324. https://doi.org/10.1017/S1366728912000466

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Trofimovich, P., & Kennedy, S. (2014). Interactive alignment between bilingual interlocutors: Evidence from two information-exchange tasks. *Bilingualism: Language and Cognition*, *17*(04), 822–836. https://doi.org/10.1017/S1366728913000801

van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language*, *92*, 26–42. https://doi.org/10.1016/j.jml.2016.05.003

Van Berkum, J. J. A. (2010). The brain is a prediction machine that cares about good and bad—any implications for neuropragmatics. *Italian Journal of Linguistics*, *22*(1), 181–208.

van Berkum, J. J. A. (2013). Anticipating communication. *Theoretical Linguistics*, *39*(1–2). https://doi.org/10.1515/tl-2013-0004

van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early Referential Context Effects in Sentence Processing: Evidence from Event-Related Brain Potentials. *Journal of Memory and Language*, *41*(2), 147–182. https://doi.org/10.1006/jmla.1999.2641

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating Upcoming Words in Discourse: Evidence From ERPs and Reading Times.

Journal of Experimental Psychology: Learning, Memory, and Cognition, *31*(3), 443–467. https://doi.org/10.1037/0278-7393.31.3.443

Van Hell, J. G., & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, *1*(3), 193–211. https://doi.org/10.1017/S1366728998000352

Weber, A., & Broersma, M. E. (2012). Spoken word recognition in second language acquisition. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Hoboken, NJ, USA: John Wiley & Sons, Inc. Retrieved from http://doi.wiley.com/10.1002/9781405198431

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1–25. https://doi.org/10.1016/S0749-596X(03)00105-0

Weber, A., & Scharenborg, O. (2012). Models of spoken-word recognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(3), 387–401. https://doi.org/10.1002/wcs.1178

Wicha, N. Y. Y., Bates, E. A., Moreno, E. M., & Kutas, M. (2003). Potato not Pope: human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters*, *346*(3), 165–168.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating Words and Their Gender: An Event-related Brain Potential Study of Semantic Integration, Gender Expectancy, and Gender Agreement in Spanish Sentence Reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272–1288. https://doi.org/10.1162/0898929041920487

Woumans, E., Ceuleers, E., Van der Linden, L., Szmalec, A., & Duyck, W. (2015). Verbal and nonverbal cognitive control in bilinguals and interpreters. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *41*(5), 1579–1586. https://doi.org/10.1037/xlm0000107

Zirnstein, M., van Hell, J. G., & Kroll, J. F. (2018). Cognitive control ability mediates prediction costs in monolinguals and bilinguals. *Cognition*, *176*, 87–106. https://doi.org/10.1016/j.cognition.2018.03.001

CHAPTER 2
PREDICTING UPCOMING INFORMATION IN NATIVE-LANGUAGE AND NON-NATIVE-LANGUAGE AUDITORY WORD RECOGNITION[1]

Monolingual listeners continuously predict upcoming information. Here, we tested whether predictive language processing occurs to the same extent when bilinguals listen to their native language vs. a non-native language. Additionally, we tested whether bilinguals use prediction to the same extent as monolinguals. Dutch-English bilinguals and English monolinguals listened to constraining and neutral sentences in Dutch (bilinguals only) and in English, and viewed target and distractor pictures on a display while their eye movements were measured. There was a bias of fixations towards the target object in the constraining condition, relative to the neutral condition, before information from the target word could affect fixations. This prediction effect occurred to the same extent in native processing by bilinguals and monolinguals, but also in non-native processing. This indicates that unbalanced, proficient bilinguals can quickly use semantic information during listening to predict upcoming referents to the same extent in both of their languages.

[1] Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition,* 20(05), 917–930. https://doi.org/10.1017/S1366728916000547

INTRODUCTION

In monolingual (native) language comprehension, people continuously generate predictions about upcoming input (e.g., Altmann & Kamide, 1999; Boland, 2005; DeLong, Urbach, & Kutas, 2005). In a seminal paper, Altmann and Kamide (1999) studied prediction in auditory language comprehension using a visual world paradigm. Participants listened to sentences such as *The boy will eat the cake* or *The boy will move the cake*. Eye movements were recorded while participants viewed a visual scene with four objects that could all be moved, but in which only one object (the cake) was edible. When participants heard the verb *eat*, participants initiated fixations to the picture of the cake more often before the onset of the word *cake* than after hearing the verb *move*. Altmann and Kamide concluded that the sentence context pre-activated the representation of the target word. Various recent models of monolingual sentence comprehension have now incorporated predictive processing (e.g., Levy, 2008; MacDonald, 2013; Pickering & Garrod, 2013).

Using context information to generate predictions is fundamental in efficient language processing: It can speed up processing, solve ambiguities, and help the listener determine when to start an overt response in a dialogue (Kutas, DeLong, & Smith, 2011; Van Berkum, 2010). These facilitatory functions could be particularly relevant in L2 comprehension, which is often considered to be slower, less accurate, and more resource-consuming than L1 processing (Cook, 1997; Hahne, 2001; Weber & Broersma, 2012). On the other hand, L2 processing difficulty may also impede efficient prediction during language comprehension. However, in spite of its possible increased importance, there is very little research about whether bilinguals predict input in their L2 like native speakers do in L1 or whether L2 words and their features are just integrated incrementally when they are encountered in the input rather than before.

In a recent review, Kaan (2014) suggested that predictive processing in L2 is not inherently different from predictive processing in L1, but it may be modulated by factors associated with non-native comprehension. For example, it is often assumed that predictions are based on statistical regularities extracted from the input throughout a person's life time (e.g., Bar, 2007; MacDonald, 2013). However, information stored in memory about how often a word tends to occur in a certain context (e.g. an edible object following the verb *eat*) may be different in L2 speakers than in L1 speakers because the L2 has usually been practiced less (Gollan, Montoya, Cera, & Sandoval, 2008) and in different settings (e.g. native learning versus classroom learning). Less or different input in L2 may affect the content and strength of predictions. Importantly, if L2 is practiced less than L1, representations of lexical form, meaning and use as well as the links between them may be less consistent and less accurate in L2 (Gollan et al., 2008). Weaker representations may lead to less efficient retrieval. And less efficient retrieval of lexical form or semantic associations may in turn lead to slower, less accurate or weaker predictions. Likewise, because bilinguals divide language use between L1 and L2, and therefore also have less L1 practice, L1 processing too may be different for monolinguals and bilinguals. If inconsistency of lexical representations indeed affects prediction skill during comprehension, then prediction skill is expected to increase with increased consistency of representations. This implies that predictive processing in L2 should become more native-like as L2 proficiency increases.

Furthermore, lexical competition is increased in L2 processing because of simultaneous activation of L1 words and because L2 speakers often misperceive phonemes, thereby increasing the number of words perceived as similar (Lagrou, Hartsuiker, & Duyck, 2013a; Weber & Cutler, 2004). Increased competition can cause a delay in the selection of a predicted word, as well as in processing the context information used to generate a prediction. Finally, a number of other factors are thought to modulate prediction in monolingual language processing, such as resource limitations, emotional state and cognitive control. Kaan (2014) suggests that the effect of each of these factors may in turn interact with processing language (native or non-native), so that L2 data is required to evaluate the generalizability of each demonstration of prediction in monolingual language processing.

Some studies reveal effects of semantic context on target word recognition (Chambers & Cooke, 2009; FitzPatrick & Indefrey, 2007; Lagrou, Hartsuiker, & Duyck, 2013b) in L2 processing. However, effects found at presentation of the target word do not allow to distinguish facilitation of semantic integration from semantic prediction. A constraining sentence context may facilitate word integration upon presentation of the word in L2 processing, but whether or not bilinguals actively predict information, online and during sentence processing, to the same extent in L1 and L2, remains unclear.

Prediction in L2 Reading

In a study in the visual domain by Martin et al. (2013), native speakers of English and late Spanish-English bilinguals read sentences in English with predictable or less predictable sentence-final nouns. Event-related potentials were measured at the article preceding the sentence-final noun. The article was always congruent with the final noun, but not with the expected noun (e.g. *Since it is raining, it is better to go out with* <u>*an*</u> *umbrella [EXPECTED]/* <u>*a*</u> *raincoat [UNEXPECTED]*). If participants indeed predicted *umbrella*, a semantic anomaly effect should be elicited by the article *a* relative to *an*, because *a* is incongruent with *umbrella*. Thus, the target for prediction is the lexical form and the congruent article. The target is predicted based on semantic information from the sentence context. Martin et al. indeed found an N400-effect for the incongruent article for L1 readers, but not for L2 readers. The lack of an effect on the article was taken to indicate that L2 readers did not predict the target word (at least not as efficiently as L1 readers). For the target noun, the authors did find a significant N400-effect in central and parietal regions in both L1 and L2 readers, but the effect was significantly larger in L1 than in L2 readers. The N400-effect on the noun showed that even though the participants reading in L2 did not predict upcoming input, integration of a target word in the sentence was still easier if the sentence was constraining.

The lack of a prediction effect on the article in L2 comprehension in Martin et al. 's study (2013) may have resulted from the particular manipulation used. In particular, the lexical prediction effect was measured on the basis of the congruency of an article (*a/an*) with the predicted word. The particular phonological agreement rule manipulated does not exist in the bilingual participants' L1. Martin et al. (2013) tested whether a group of intermediate L2 proficient participants, not participating in their experiment, knew the phonological article-noun agreement rule. Both an online and an offline test showed that intermediate L2 proficient participants were sensitive to the agreement rule. However, the intermediate L2 proficient group actually participating in the experiment may not have been able to apply the rule quickly enough for a prediction-incongruent determiner to modulate the N400 effect. Therefore, in a second study in the visual domain, Foucart, Martin, Moreno and Costa (2014), used a similar sentence reading paradigm but measured the prediction effect by manipulating prediction congruency of the determiners' gender in Spanish sentences (e.g. *The pirate had the secret map, but he never found the [masc] treasure [EXPECTED]/ the [fem] cave [UNEXPECTED] he was looking for.*). As in Martin et al., the target for prediction is the lexical form and the congruent article. The target is predicted based on semantic information from the sentence context. However, in this study the gender agreement rule between the target article and noun existed both in the late bilingual participants' L1 (French) and L2 (Spanish). Here, the authors found an effect of congruency of the article and the predicted noun on the N400 elicited by the article both in L1 reading (by Spanish monolinguals and early Spanish-Catalan bilinguals) and in L2 reading (by late French-Spanish bilinguals), although the effect lasted for a shorter time in the late bilingual group. The results demonstrate that bilinguals reading in L2 can use semantic information from the sentence context to predict upcoming words and their gender. Foucart et al. suggested that the similarity between the article-noun agreement rule in late bilingual participants' L1 and L2 may have made it easier for the participants to generate a prediction in time. In addition, half of the expected nouns included in the experiment were cognates, possibly adding to the facilitatory effect. The two studies described above show that bilinguals can predict lexical information in sentence reading, but that whether or not prediction occurs may depend on L1 and L2 language similarity.

Prediction in L2 Listening

Both studies described above were conducted in the visual domain, but predictive language processing may well be more challenging in the auditory modality. For instance, the fact that auditory input unfolds over time, unlike written input, may make prediction more relevant because the listener cannot return to prior input or influence input rate, unlike reading [2]. Predictive processing may also be more difficult in the auditory modality than in the visual modality for bilinguals because of increased cross-language co-activation due to misperceptions and misrepresentation in listening (Weber & Cutler, 2004).

Foucart, Ruiz-Tada and Costa (2015) tested prediction in the auditory modality using an EEG paradigm similar to Foucart et al. (2014). Again, the target for prediction was the lexical form with the congruent article, and predictions were based on semantic information from the sentence context. The authors found that bilinguals listening in L2 are able to predict upcoming words based on sentence context. The participants in this study were all bilingual and they were only tested in their L2. Therefore, no direct comparison could be made between the size of the effect in L1 and L2 in bilinguals, or between the size of the effect in monolinguals (L1) and bilinguals (L1 or L2).

Visual world paradigm studies on prediction in L2 auditory processing have mainly focused on prediction based on morpho-syntactic information. In a visual world experiment, Hopp (2013) investigated whether German native and English-German bilingual listeners would show predictive looks to target objects whose gender agreed with an article in the auditory signal. Like native listeners, English-German bilinguals listening in L2 were more likely to look at the target objects whose gender agreed with an afore-mentioned article before the onset of the target object in the auditory signal, but only in the bilinguals who had native-like mastery of gender assignment.

Hopp (2015) used a visual world paradigm to investigate whether English-German bilinguals integrate morphosyntactic and verb semantics information to generate predictions about upcoming

[2] In Martin et al. (2013), the first half of each stimulus sentence was presented on the screen as a whole. After pressing spacebar, one word was presented every 700 ms.

semantic input during L2 auditory comprehension. In this experiment, picture displays including three possible actors and a control object were paired with an SVO (e.g. *The$_{NOM}$ wolf kills soon the$_{ACC}$ deer*) or an OVS (e.g. *The$_{ACC}$ wolf kills soon the$_{NOM}$ hunter*) sentence in German. Native listeners were more likely to look at expected patients (*the deer)* before the onset of the second NP in SVO sentences and at expected agents (*the hunter*) in OVS sentences. English-German bilinguals on the other hand, were more likely to fixate patients before the onset of the second NP, independently of the case marking (nominative or accusative) of the first NP. Hopp concluded that there was an effect of semantic prediction in L2 based on information extracted at the verb, but that case information did not modulate predictions like in L1 listeners. Bilingual participants seemed unable to apply an L2 agreement rule not present in their L1 on the fly, or at least not quickly enough to support prediction. Hopp's findings are in line with recent findings of Mitsugi and Macwhinney (2016), who demonstrated that L1 English learners of Japanese with good offline knowledge of the Japanese case-marking system were unable to employ this knowledge online in order to generate predictions in a visual world eye-tracking experiment.

Dussias, Valdés Kroff, Guzzardo Tamargo and Gerfen (2013) also focused on prediction based on morpho-syntactic information, specifically, prediction based on article-noun gender agreement. A group of English-Spanish bilinguals (high and low proficiency), Italian-Spanish bilinguals and Spanish monolinguals saw a display with two pictures of items with the same or different grammatical gender. While looking at the display, they heard a sentence with an article that either agreed with the gender of one of the two items in the display, or with both. Spanish monolinguals looked at the target picture sooner in the different gender condition (when the article was a cue) than in the same gender condition. Highly proficient English-Spanish bilinguals, but not low proficient English-Spanish bilinguals, also looked at the target picture earlier in the different gender condition. Unlike the low proficient English-Spanish bilinguals, low proficient Italian Spanish bilinguals looked at the target picture significantly earlier in the different gender condition, but only when the target item was feminine. Dussias et al. 's results suggest that highly proficient bilinguals use gender cues to anticipate information like monolinguals do, whereas low proficient bilinguals do not, unless their native language has a similar article-noun gender agreement system. Even though the effects Dussias et al. found for monolinguals and highly proficient bilinguals are likely to be anticipatory in nature, given their time course, the authors do not distinguish between effects anticipation and facilitation of integration.

These recent visual world studies on prediction in L2 listening reveal that it is especially difficult for bilinguals to process morpho-syntactic features quickly enough to use them as a cue to generate predictions in L2. However, it remained unclear whether bilinguals also have difficulty anticipating semantic information in L2 processing, which would always lead to weaker L2 prediction effects, or, whether they selectively have difficulty applying language-specific, and difficult, grammatical rules quickly enough during predictive processing. Hopp (2015) explicitly distinguishes prediction based on verb semantics and prediction based on case-marking. However, as Hopp proposes, the significant effect of prediction based on verb semantics (predictive looks to the patient object in both SVO and OVS sentences) in L2 listening can be interpreted in two ways: Either the L2 listeners used semantic information extracted at the verb to guide predictive looks towards the most plausible sentence object in the picture display (the patient), or, on the basis of the first NP, fixations were directed to a plausible patient object, regardless of verb semantics. Therefore, it remains unclear whether bilinguals are able to use verb semantics to guide their predictions during non-native sentence comprehension like they do in L1.

Koehne and Crocker (2015) provided evidence that language learners are able to use semantic restrictions at the verb to predict upcoming referents. Participants learned novel, artificial verb, subject (*man* and *woman*) and object names by exposure to verbs with visual context, followed by exposure to nouns in SVO sentence context, in a visual world paradigm. Anticipatory eye-movements to the sentence target objects were found during presentation of the constraining verb. As each verb type was combined with each subject type, the anticipatory eye-movements to the target object could not have been based on information extracted at the sentence subject alone. Koehne and Crocker show that people can use verb semantics to predict upcoming information in early language learning. However, instruction specifically stressed semantic processing of the sentences. Also, a limited number of artificial verbs (six at most) and objects (18 at most) were used in the study. These two factors may have greatly inflated predictive processing when compared to natural L2 language processing.

Present Study

All previous studies on anticipating information in L2 listening have either focused on L2 listening alone, or they have compared a group of L2 listeners to a group of L1 listeners in a between-participants design. In the present experiment, Dutch-English bilinguals were tested in the native and non-native language. In addition, an English monolingual control group was tested in order to compare L1 with L2 listening in the same language (English) and L1 (English) listening by monolinguals with L1 (Dutch) listening by bilinguals. Comparing predictive processing within participants is important, as recent studies have shown effects of cognitive factors such as verbal fluency, vocabulary size (Rommers, Meyer, & Huettig, 2015), working memory and processing speed (Huettig & Janse, 2016) on predictive language processing. There may also be factors inherent to bilingualism (and not L2 processing) that affect predictive processing. For example, bilinguals activate lexical information in both languages during L1 and L2 processing (e.g. Lagrou et al., 2013a). Bilinguals may therefore activate more information during language processing which in turn may slow down the prediction process. In addition, some authors suggest that bilinguals have increased cognitive control abilities compared to monolinguals (Woumans, Ceuleers, Van der Linden, Szmalec, & Duyck, 2015). Increased cognitive control may help suppress irrelevant information during predictive processing. For example, Zirnstein, Hell and Kroll (2015) recently found that that processing costs for unverified predictions were larger in low-control than in high-control bilingual participants. In this experiment, we will compare bilinguals listening to speech in L1 and L2 to eliminate effects of individual differences. As a control experiment, we will also compare bilinguals (L2) to monolinguals (L1) listening to the same language (English). Finally, to test whether there are any effects of speaker bi- or monolingualism on predictive language processing we will compare prediction effects in L1 processing in bilinguals (Dutch) to L1 processing in monolinguals (English).

Here, a visual world paradigm based on Altmann and Kamide's (1999) task was used. Participants listened to sentences such as *Mary knits a scarf* or *Mary loses a scarf*. Eye movements were recorded while participants viewed a visual scene with four objects that could all be *lost* (neutral condition), but in which only one object (the scarf) was *knittable* (constraining condition). If participants predicted the target object in the constraining condition, this would result in a higher proportion of looks to the target object in the constraining condition than in the neutral condition before the onset of the target in the auditory stimulus. Based on Kaan (2014) we expected that bilinguals listening in L2 would not predict semantic properties of upcoming referents as fast and to the same extent as when listening in L1 because of modulating factors associated with L2 language processing, such as differences in stored statistical regularities and weaker, less accurate lexical representations. Further, we expected that bilingual participants listening in L1 would not predict semantic input to the same extent as monolinguals do in L1. This would be in line with the weaker links hypothesis (Gollan et al., 2008) of bilingual language processing. This hypothesis states that bilinguals divide language use between L1 and L2, and therefore have less practice in each of their languages. Less practice in each language should lead to weaker links between semantics and phonology in bilinguals than in monolinguals and thereby to slower lexical access. In turn, these weaker links may result in slower or weaker predictive processing.

As opposed to previous studies on predictive processing in the non-native language, we opted for a design in which no language-specific agreement rule needed to be applied by the participants on the fly in order to measure the prediction effect or in order for the participant to make a prediction. This way, if we find an attenuation of the prediction effect in non-native listening, it cannot be attributed to difficulty applying a non-native agreement rule on the fly.

Finally, previous studies have suggested that predicting upcoming information during language processing serves as a learning mechanism (Dell & Chang, 2013; Koehne & Crocker, 2015; Mani & Huettig, 2012). For example, Mani and Huettig (2012) found a significant positive correlation between prediction skill and expressive vocabulary in children. We therefore expect that prediction effects should be modulated by language proficiency, so that bilinguals with a higher proficiency score show a stronger prediction effect than bilinguals with a lower proficiency score.

METHODS

Participants

Bilinguals. Thirty native speakers of (Belgian or Netherlands) Dutch took part in the experiment (5 men and 25 women, mean age 24 years, range 20-41). They were recruited from the Ghent University participant database. All signed informed consent. All participants reported Dutch as their dominant and most proficient language in the LEAP-Q questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007), and English as their second (25 participants) or third (5 participants) language. Belgian and Dutch students typically start to learn English at age ten or eleven in school, and their English proficiency is relatively high because of regular input from popular media and study books. None of the participants had immersion experience in an English-dominant environment. On average the participants reported to be exposed to English 17% of the time, versus 73% to Dutch. Besides knowledge of English and Dutch, twenty-eight participants had knowledge of French, and nineteen participants had knowledge of German. Fewer than six participants had knowledge of other languages such as Spanish, Turkish, Portuguese, Polish or Italian. To assess language proficiency in both languages, participants carried out the LexTALE vocabulary knowledge test (Lemhöfer & Broersma, 2012) and provided self-ratings. The LexTALE is an unspeeded 60-item lexical decision task. It is an indicator of word knowledge and general language proficiency (Lemhöfer & Broersma, 2012). The bilinguals' mean LexTALE scores and self-ratings are reported in Table 1. The LexTALE score and self-ratings show that the bilingual participants were more proficient in their native (Dutch) than in their non-native language (English).

Monlinguals. Thirty monolingual native speakers of English participated in the experiment (4 men and 26 women, mean age 20 years, range 18-28). They were recruited from the Southampton university participant database. All signed informed consent. The monolinguals' mean LexTALE scores and self-ratings are reported in Table 1. The LexTALE score shows that the bilingual and monolingual participants were matched on L1 proficiency.

Table 1. *Participants' Mean Scores on Proficiency Tests and Mean ratings.*

	Bilinguals Dutch (SD)	Bilinguals English (SD)	Monolinguals English (SD)	p-value[c] bilinguals L1 vs. L2	p-value[d] monolinguals L1 vs. bilinguals L1	p-value[e] Monolinguals L1 vs. bilinguals L2
Lextale[a]	86.13 (5.54)	78.50 (10.49)	87.83(7.97)	<0.001	.34	<0.001
Rating speaking	9.2 (0.75)	7.3 (1.34)	9.6 (0.72)	<0.001	.03	<0.001
Rating listening	9.3 (0.79)	8.1 (0.73)	9.5 (0.78)	<0.001	.07	<0.001
Rating reading	9.3 (0.66)	8.0 (1.22)	9.3 (0.92)	<0.001	.44	<0.001
Mean ratings [b]	9.3 (0.7)	7.8 (0.9)	9.5 (0.71)	<0.001	.09	<0.001

[a] Scores consist of percentage correct, corrected for unequal proportion of words and nonwords (Lemhöfer & Broersma, 2012).

[b] Score based on means of self-assessed ratings on a scale of 1 to 10 (1=very low, 10=perfect) of speaking, listening and reading.

[c] Reported p-values indicate significance levels of dependent samples t-tests between scores for Dutch and English in bilinguals. Df of all t-tests= 29.

[d] Reported p-values indicate significance levels of independent samples t-tests between scores for bilinguals in Dutch and monolinguals in English. Df of all t-tests= 29.

[e] Reported p-values indicate significance levels of independent samples t-tests between scores for bilinguals and monolinguals in English. Df of all t-tests= 29.

Materials and Design

Eighteen stimulus sets were created. Each set consisted of a four-picture display, two sentences in Dutch and their translation equivalents in English. One of the two sentences was constraining; the other sentence was neutral. In the constraining condition, only one of the objects in the display was appropriate after the verb, whereas all objects in the display were appropriate after the verb in the neutral condition (see Figure 1). Appendix 2A contains the constraining and neutral verbs as well as the objects in the display for each stimulus set[3].

Figure 1. Example Picture Display. The sentences belonging to this display were: *Mary reads a letter* and *Mary steals a letter*.

Likewise, eighteen filler sets were created. Each set again consisted of a display with four pictures, two sentences in Dutch and their English translation equivalents. In the filler sets, sentences could apply to either no, or two or three objects in the display. The stimulus and filler

[3] To check whether the semantic association strength between the verb and the target picture name was stronger in the constraining than in the neutral condition, and whether the association strength was similar across languages, we obtained a measure for semantic association from the snaut tool (Mandera, Keuleers, & Brysbaert, in press). In snaut, the association strength between verb and target is calculated based on co-occurrences in large text corpora. The stronger the association strength, the lower the measure.As expected, paired t-tests pointed out that there was a stronger semantic association between the verbs and targets in the constraining condition than in the neutral condition ($p<.001$ for Dutch and $p=.002$ for English). Also, there was no significant difference between the association strengths in our English and Dutch stimuli ($p=.18$), indicating that our stimuli sentences were matched for semantic association strength between languages.

sentences were randomly assigned to two stimulus lists with the constraints that two sentences belonging to the same set were never in the same list, and each list contained an equal number of neutral and constraining sentences.

Pictures. The pictures were line drawings from a normed database by Severens, Van Lommel, Ratinckx, and Hartsuiker (2005). Each target picture was included as unrelated picture in another stimulus set. This way, we ensured that target pictures did not inherently draw more overt visual attention than unrelated pictures. The names of the objects in each display were never semantically associated with the verb in the neutral condition and only the target object could be associated with the verb in the constraining condition (association norms from Deyne, Navarro, & Storms, 2013). The onsets of the names of objects in one display were never identical, nor were they identical to the onsets of the accompanying verbs.

Three repeated-measures ANOVAs with language (native, non-native) and picture type (target, distractor) as factors showed that object names were matched for frequency, phoneme count, and syllable count across languages and conditions (ps> .10) (Table 2). The selected object names were orthographically dissimilar (normalized orthographic Levenshtein distance $\leq .50$, $M=.15$, $SD=.13$[4]). The pictures had a mean H-statistic (a name agreement index) in Dutch of .62 ($SD=.49$) (Severens et al., 2005)[5]. To our knowledge, no name agreement scores are available for the picture set for bilinguals in L2.

Sentences. Simple four-word SVO sentences were constructed for this experiment. The subject of the sentence was kept constant across all trials (*Mary* in English, *Marie* in Dutch). Repeated measures ANOVAs with language (native vs. non-native) and condition (neutral vs. constraining) as factors showed that verb frequency, phoneme count, and syllable count were matched across languages and conditions (all ps$\geq .10$). Table 2 reports the lexical characteristics of the stimuli in English and in Dutch.

[4] 0=no overlap, 1=identical (Schepens, Dijkstra, Grootjen, & van Heuven, 2013).
[5] The mean H-statistic of the full picture set of Severens et al. (2005) was 1.00 with scores ranging from 0 to 3.19. Lower H-statistic scores indicated higher name agreement.

Table 2

Mean Lexical Characteristics of Dutch (native) and English (non-native) Stimuli.

		Frequency[a]	Phoneme Count[b]	Syllable Count[b]
Picture name Dutch	Target	4.39 (.61)	4.17(1.54)	1.44(.71)
	Distractor	4.28 (.29)	4.19(.76)	1.31(.37)
Picture name English	Target	4.46 (.58)	4.17 (1.20)	1.44 (.62)
	Distractor	4.29 (.27)	4.15 (.68)	1.48 (.26)
Verb Dutch	Neutral	3.85(.60)	5.28 (1.56)	1.44(.62)
	Constraining	3.48 (.77)	4.83(1.04)	1.44(.51)
Verb English	Neutral	3.74 (.42)	4.78(.94)	1.22(.43)
	Constraining	3.50 (.62)	4.78(.88)	1.33(.49)

Note. Standard deviations are indicated in parentheses.

[a] Zipf value (log10(frequency per million*1000) (van Heuven, Mandera, Keuleers, & Brysbaert, 2014) retrieved from the SUBTLEX-US and SUBTLEX-NL databases (Brysbaert & New, 2009; Keuleers, Brysbaert, & New, 2010)

[b]. CELEX database (Baayen, Piepenbrock, & Gulikers, 1995).

The article preceding the sentence final noun was always indefinite, and English nouns never started with a vowel. This ensured that the article could not be used as a prediction cue.

Recordings. Sentences were recorded in a sound attenuating room. A female native speaker of Dutch (34 years old) who majored in Dutch and English linguistics and literature at university pronounced the sentences for both the English and the Dutch recordings. English monolinguals rated her accent as 5.3 on a scale from 1 (very foreign accent) to 7 (native accent). We chose this speaker for our study because of her clear pronunciation in Dutch and English, and experience in recording psycholinguistic stimuli. Each sentence was recorded three times; the recording that we judged to have the most neutral prosody was selected for the experiment.

The length of the recording frames starting at verb offset, and ending at noun onset initially differed significantly between Dutch and English ($t(35)=10.87$, $p<.001$). In the non-native condition, participants would therefore have less time to generate predictions about upcoming referents than in the native condition. To eliminate this confound, the fragment was lengthened by a factor 1.2 for the English sentences and shortened by a factor 0.8 for the Dutch sentences, using Praat (Broersma & Weenink, 2014). This way, the length of the recording fragments was matched across languages (*p*s $\geq$.10). The mean length of the verb onset – noun onset frame was now 691 ms in Dutch and 708 ms in English. None of the participants indicated having noticed the manipulation of the auditory stimuli.

Procedure

Participants were seated at a comfortable distance from the screen. They received written and verbal instructions to listen carefully to the sentences and to look at whatever they wanted as long as their gaze would not leave the screen (Huettig & Altmann, 2005; McQueen & Huettig, 2012). There was no explicit task. Eye movements were recorded from the right eye with an Eyelink 1000 eye-tracker (SR Research) with a sampling frequency of 1000 Hz. After successful calibration, the experiment began with two practice trials.

A fixation cross appeared on the screen for 500 ms, followed by the presentation of the four pictures in a two-by-two grid on the screen. Picture location was randomized. The auditory stimulus started to play 2200 ms after picture onset. This time lag was included to ensure that participants had enough time to see every object on the screen before verb onset. The trial ended when the sentence finished, and the next trial was started by the experimenter after drift correction. Bilingual participants were presented with the stimuli in one of the lists in a Dutch (native) block and with the other list in an English (non-native) block. Language and list order were counterbalanced. Monolingual participants were presented with the stimuli of one list in the first block and with the stimuli of the other list in the second block. Both lists were presented to the monolinguals in English. List order was counterbalanced. In each block, the participants heard nine constraining and nine neutral sentences. Across the two blocks, none of the verbs were repeated, but the object displays were repeated. The eye tracker was recalibrated between the two blocks. The entire experiment took approximately 17 minutes.

After the experiment, participants completed the following additional tests: LexTALE Dutch, LexTALE English (Lemhöfer & Broersma, 2012) (see Table 1 for results), backward translation of the English verbs used in the experiment (bilinguals only), backward translation of the English nouns used in the experiment (bilinguals only), and the LEAP-Q language background questionnaire(Marian et al., 2007). The tests were presented in that order on a Macbook in a quiet room. Completion of the additional tests took approximately 25 minutes.

RESULTS

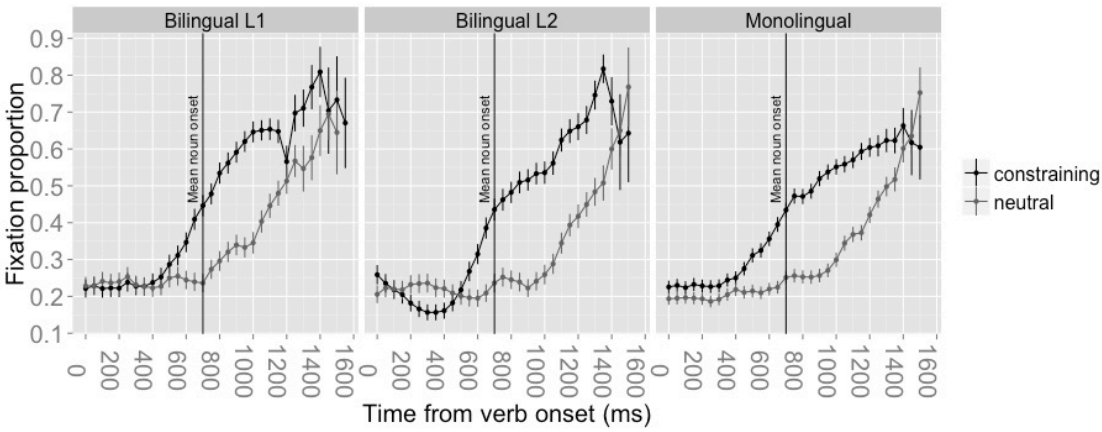


Figure 2 shows the time-course of target fixation as a function of condition for each language and speaker group. These probabilities reflect the number of samples of eye-data within a 50 ms time bin in which there was a fixation on the target picture, averaged over subjects and items.

Figure 2. Results. Time course of fixation probability to target by language (native, non-native) and condition (constraining, neutral) starting from verb onset. Note: The mean noun onset is aligned to the 50ms bin within which they fall. Whiskers indicate the mean ± standard error.

The graph shows that participants were more likely to fixate on target objects in the constraining condition than in the neutral condition. Fixation proportions for the constraining and neutral conditions start to diverge well before the mean noun onset time in each of the three groups.

The starting point of the time frame for our analysis was chosen based on visual inspection of a plot of the time-course of the grand mean of fixation probability (over languages and listener types) and was defined as the first 50 ms bin after verb onset in which the grand mean fixation probability began a rising trend (Barr, 2008). This method is conservative because by using the grand mean the choice can not be biased by any hypothesis (Barr, 2008). As it takes approximately 200 ms to plan and execute a saccade (e.g., Matin, Shao, & Boff, 1993; Saslow, 1967), we can assume that fixations that started earlier than 200 ms after noun onset were anticipatory in nature. Thus, the time frame for the analysis started at 350 ms after verb onset and ended 200 ms after noun onset. Each trial's individual verb onset and noun onset times were used to select the data. In addition to the analysis of the full time frame we analysed the data of the first four hundred milliseconds of data in the analysis frame aggregated into 100 ms time bins. This way, we tested when the effect of condition became significant in each group. In 3.39% of the samples in time frame from verb onset until 200 ms after noun onset there was a blink and 0.17% percent of the samples were out-of-screen. The out-of-screen and blink samples were included in the total sample count used to calculate proportions of looks to the target image.

The proportions of samples in the analysis time-frame in which there was a fixation to the target image were transformed using the empirical logit formula (Barr, 2008). Our data set was analyzed with linear mixed effects models with the lme4 (version 1.1-8), car (2.0-25) and lmerTest (version 2.0-25) package of R (3.2.1) (R Core Team, 2013). This allowed for inclusion of participant, sentence and target image as random factors (Baayen, Davidson, & Bates, 2008).

For the analyses between languages in bilinguals, the fixed experimental factors were condition (constraining or neutral) and language (Dutch or English). The control variables list (A or B) and block (1 or 2) were also included as fixed factors. The models included random intercepts for participant, sentence and target picture. In each analysis we first fitted a model including all the fixed factors and interactions as well as the random intercepts for participant, sentence and target picture. If there was a significant effect of a factor, we added that factor as random slope for participant, sentence and target picture. For the comparison between listener types (monolinguals and bilinguals) in English and in Dutch listening, the fixed factors were condition (constraining or neutral) and listener type (monolingual or bilingual). All other factors were the same as in the within participants analysis[6]. To test whether there were any effects of English proficiency on predictive processing we compared each model without the factor lexTALE score (English) to the model with the factor lexTALE score and LexTALE as random slope for sentence and target picture using a likelihood ratio test. Eighteen trials were removed from the dataset because the verb was not translated correctly in the translation task that was performed after the main task, by that particular participant.

[6] In addition to our main analysis with the dichotomous factor 'condition' (neutral versus constraining), we checked whether there was an effect of the semantic association strength between verb and target pairs on fixation proportion to target images in the analysis time frame. The measure 'semantic association strength' was obtained from snaut (Mandera, Keuleers, & Brysbaert, in press) (see Footnote 2). We tested this for each analysis separately: within bilinguals (L1 and L2), between listener types (monolinguals and bilinguals) in English, and between listener types in L1 (Dutch for bilinguals and English for monolinguals). In the within-bilinguals (bilinguals in L1 and L2) analysis, there was a marginally significant effect of association strength $\beta = -3.44$, $SE = 1.64$, $t = -2.09$, $p = .056$. The stronger the association strength, the more fixations to the target image in the analysis time frame. In the between listener type analysis (English in monolinguals and bilinguals), the main effect of association strength reached significance: $\beta = -3.37$, $SE = 1.49$, $t = -2.27$, $p = .032$. Finally, in the within L1 analysis (Dutch in bilinguals, English in monolinguals), no significant effect of association strength was found ($\beta = -2.03$, $SE = 1.38$, $t = -1.47$, $p = .14$). Importantly, there were no significant interactions between association strength and language or listener type in any of the analyses. The analyses suggest that stronger semantic association yields stronger prediction. We currently have no theory as to why the effect of association strength on target fixations did not reach significance in the within L1 analysis.

Comparison within Bilinguals (L1 vs. L2)

The fixation proportion was significantly higher in the constraining condition than in the neutral condition (β = -0.54, *SE* = .12, *t*= -4.49, *p*<.001), confirming our prediction manipulation. There was no significant interaction between language (L1 vs. L2) and condition (constraining vs. neutral) (β = 0.04, *SE* = .10, *t*= .40, *p*=.69). Nor were there any other significant main effects[7]. English proficiency (lexTALE) score did not significantly improve the model fit (χ^2(19)=15.2, *p*=.71)[8].

Separate analyses for each language revealed that the effect of condition was significant in L1 (β =-0.65, *SE* = 0.17 *t*=-3.86, *p*=.001)[9], and also in L2 (β =-0.56, *SE* =0.15, *t*= -3.58, *p*<.001).

Comparison between L1 Monolingual Listening (English) and L2 Bilingual Listening (English)

The fixation proportion was significantly higher in the constraining condition than in the neutral condition (β = -0.69, *SE* = .12, *t*= -5.76, *p*<.001). The effect of condition did not interact with listener type (monolingual versus bilingual) (β = -.11, *SE* = .11, *t*= -.93, *p*=.36). Nor were there any other significant main effects. English proficiency (lexTALE) did not significantly improve the model fit (χ^2(22)=24.72, *p*=.32). The effect of condition was also significant in the data of the monolinguals only (β = -.79, *SE* = .16, *t*= -4.87, *p*<.001).

Comparison between L1 Monolingual Listening (English) and L1 Bilingual Listening (Dutch)

The fixation proportion was significantly higher in the constraining condition than in the neutral condition (β =-.72, *SE* = .13, *t*= -5.57, *p*<.0001). There was no significant interaction between listener type (monolingual vs. bilingual) and condition (β =-.07, *SE* = .12, *t*=-.61, *p*=.55). Proficiency (English LexTALE score) did not contribute significantly to the model fit (χ^2(22)= 29.21, *p*=.14).

Time Course Analyses

---

7 After processing L2, processing in L1 tends to be slowed down (Misra, Guo, Bobb, & Kroll, 2012). Therefore, anticipatory effects were expected to be smaller or start later in an L1 block following an L2 block than vice versa. Our analyses showed that the order of language blocks did not interact with the effect of prediction, neither in the analysis of the entire time frame nor in the time course analyses.

8 The result is reported for the model with LexTALE as random slope for target picture, but not for sentence. Condition was also included as random slope for participant and target picture. This was the maximum random effect structure justified by our sample (including LexTALE as random slope for sentence resulted in a model convergence error).

9 The result is reported for the model with condition as random slope for participant, but not for target image. This was the maximum random effect structure justified by our sample.

In the bilinguals, the effect of condition became significant in the third time bin of the analysis time frame (550-650 ms) (β = -.45, *SE* = .15, *t*= -2.94, *p*=.007). There was no significant interaction between language and condition (β =-.03, *SE* = .12, *t*= -0.21, *p*=.84). In a separate analysis of the bilingual data for each language, the effect of condition also became significant in the third time bin of the analysis frame in English (550-650 ms after verb onset) (β = -.47, *SE* = .20, *t*= -2.32, *p*=.03) and in Dutch (β = -.43, *SE* = .19, *t*= -2.24, *p*=.03).

In the comparison between listener types in English (L1 monolinguals vs. L2 bilinguals) the main effect of condition was not yet significant in the first two time bins (350-450ms after verb onset: β = .06, *SE* = .13, *t*= .46, *p*=.65, 450-550 ms after verb onset: β = -.19, *SE* = .14, *t*= -1.35, *p*=.18). However, the interaction between listener type and condition was significant in the first bin (β = -.25, *SE* = .12, *t*= -2.09, *p*=.04)[10], and marginally significant in the second bin (β = -.22, *SE* = .12, *t*= -1.89, *p*=.06). In the third time bin, the effect of condition became significant (β = -.55, *SE* = .15, *t*= -3.78, *p*<.001), and the interaction between listener type (monolingual vs. bilingual) and condition was no longer significant (β = -.06, *SE* = .11, *t*= -.56, *p*=.57).

Finally, we compared the two listener types in L1 (English in monolinguals vs. Dutch in bilinguals). The effect of condition became significant in the second time bin in the analysis frame β = -.28, *SE* = .13, *t*= -2.09, *p*=.04. The interaction between listener type and condition did not reach significance β = -.13, *SE* = .12, *t*= -1.02, *p*=.31.

In a separate analysis of the monolingual data, the effect of condition was significant for the first time in the second time bin in the analysis frame (450-550 ms after verb onset) (β = -.41, *SE* = .18, *t*= -2.29, *p*=.03). At that time, the effect was not yet significant for the bilinguals in English (L2) (β = -.04, *SE* = .17, *t*= -.21, *p*=.83) or in Dutch (L1) (β = -.15, *SE* = .18, *t*= -.88, *p*=.38).

DISCUSSION AND CONCLUSION

This study asked whether bilinguals predict information about upcoming referents on the basis of semantic context information during non-native comprehension, like monolinguals do in L1 comprehension. Following monolingual studies (e.g., Altmann & Kamide, 1999), we found that bilinguals use linguistic context information to generate predictions about upcoming referents in

[10] The result is reported for the model with condition as random slope for participant and target image. Listener type was included as random slope for sentence, but not for target image. This was the maximum random effect structure justified by our sample.

their non-native language (English). This effect was of comparable magnitude in L1 listening in the same participants (Dutch) and in L1 listening by monolinguals in the same language (English). In addition, bilinguals listening in L1 (Dutch) predicted upcoming semantic information to a similar extent as monolinguals listening in L1 (English). English proficiency (lexTALE score) did not affect the prediction process. These findings confirm that bilinguals listening to non-native input are able to rapidly integrate auditory and visual input to constrain the subsequent domain of reference[11]. Consistent with the weaker links hypothesis (Gollan et al., 2008), time-course analyses suggested that bilinguals listening in either L1 or L2 predicted upcoming information slightly slower than monolinguals.

Kaan (2014) argued that predictive processing in a non-native language is not inherently different from predictive processing in the native language, but that other factors associated with non-native processing (e.g. cross-linguistic competition, inconsistent lexical representations in L2) can modulate prediction. No modulation of the prediction effect in the non-native language was found in the present study. Perhaps the modulating factors Kaan discussed only play a role under specific circumstances such as in sentences with infrequent words or cognates. Infrequent words are likely to have inconsistent representations because they are practiced less often. Also, no large cross-linguistic interference effects were expected because target words were never cognates. Furthermore, in visual world paradigm experiments like the present one, prediction processes may be facilitated as compared to EEG studies (e.g., Foucart et al., 2015), because visual candidates for prediction (pictures) are provided with each sentence (Kamide, 2008). Target words or target semantics were likely to be pre-activated along with the three other candidates.

Like us, Foucart et al. (2015) found a significant prediction effect in L2 speech processing using an EEG paradigm. The authors measured the modulation of the N400 effect elicited by an article that was gender congruent or incongruent with the predicted noun in L2 listening. The article-noun agreement rule manipulated in this experiment exists both in the bilingual participants' L1 and L2. Foucart et al. therefore suggested that prediction can be accomplished in L2 processing if the L2 is similar to the L1. Unlike in Foucart et al.'s study, no cognates were included as target words in our visual world experiment. Therefore, the prediction effect found in non-native listening in our experiment did not depend on target similarity between languages. However, English and

[11] Note that in the current design it is possible that sometimes the target picture (visible before the onset of the auditory signal) primed the verb, because of a strong semantic association between verb and target picture. This could strengthen the further prediction. It is impossible to dissociate the effect of associative strength between verb and target on verb priming vs. target prediction. As association strength did not differ between languages or listener groups (see Footnote 2), our conclusions still stand.

Dutch are typologically similar languages, therefore Foucart et al.'s suggestion that prediction in L2 is facilitated by L1 and L2 similarity is still viable. The present experiment complements Foucart et al.'s results because we make a direct comparison between the prediction effect in bilinguals in L1, in L2 and in monolinguals in L1, and show that the magnitude of the prediction effect is the same in each language and speaker group.

Hopp (2015) also found an effect of prediction in non-native listening using a visual world paradigm. Unlike native listeners whose predictions were based on semantic and case-marking information, the non-native listeners were unable to use case-marking information to modulate predictions. Non-native listeners' predictive looks to likely patient objects may have been based on the semantic information extracted at the first NP in the sentence regardless of verb semantics, or, on a combination of semantic information of the first NP and verb semantics. In the present experiment no picture of the first NP in the sentence was shown in the display, and only the verb distinguished the neutral from the constraining condition. Therefore, this study confirms that bilinguals listening in L2 can use verb semantics in order to predict features of upcoming input to the same extent in L1 and L2.

Previous studies showed that bilinguals have difficulty with predicting L2 input based on morphosyntactic information such as case or gender information (Dussias et al., 2013; Hopp, 2013, 2015). Predicting upcoming words together with morphosyntactic information (the gender of an article) is also difficult for bilinguals (Martin et al., 2013), unless the second language shares morphosyntactic features (e.g. gender-noun agreement rules) with the first (Foucart et al., 2014, 2015). However, in line with Koehne and Crocker (2015) the results of the present study show that bilinguals have no difficulty predicting input based on semantic information. This suggests that bilinguals predict to a similar extent in L2 as monolinguals do in L1, but that problems arise only when morphosyntax is involved, perhaps because of difficulty applying morphosyntactic agreement rules online quickly enough. An interesting question for future research would be whether increased processing speed (e.g. increased speech rate) would lead to difficulty using semantic information to generate predictions in L2 as well.

Speaker accent can affect speech processing (Adank, Evans, Stuart-Smith, & Scott, 2009; Lagrou et al., 2013b; Weber, Betta, & McQueen, 2014). Dutch-English bilinguals in Belgium are frequently exposed to non-native speakers in school and work settings, and in the media. Therefore, they are familiar with Dutch-accented English like the accent of the speaker in the experiment. A previous study from our lab (Lagrou et al., 2013b) showed that in a lexical decision task, Dutch-

English bilinguals responded faster to English stimuli pronounced by a native speaker than to English stimuli pronounced by a non-native speaker. If words are recognized more slowly by L2 listeners when pronounced by an L2 speaker than by an L1 speaker (Lagrou et al.), then an interaction effect of language (L1 or L2) with prediction of upcoming information is likely to be more pronounced when the speaker of the experimental stimuli is a non-native speaker. No such interaction was found in the present experiment. Whether various strengths of non-native accents affect the prediction process differently in L1 and L2 listeners remains an open issue[12].

English proficiency as measured with LexTALE (Lemhöfer & Broersma, 2012) did not affect the magnitude of the prediction effect in bilinguals and monolinguals. This may be due to the high level of proficiency of our participants, although these were still clearly unbalanced bilinguals who only use L2 during a small proportion of their time. Alternatively, there may not have been sufficient variance (bilinguals M=78.5, SD=10.49) to detect an interaction effect of proficiency with prediction skill. Conversely, production skill and not recognition skill may be an indicator of prediction skill (Mani & Huettig, 2012), and the LexTALE does not tap into production skill directly. In any case, the present data show that the proficiency level of these unbalanced bilinguals suffices for predictive language processing similar to that in the native language.

The time course analyses showed that prediction effects reached significance 100 ms later for bilinguals (in both languages) than for monolinguals. One theoretically interesting interpretation would be that activation and prediction develops slower for bilinguals. However, this may also merely be due to lack of power in the bilingual data sets. The monolinguals were exposed to both stimuli lists in English whereas the bilinguals were exposed to one list in each language. Therefore, the monolingual data set is twice the size of the bilingual data sets of each language, which increases power to detect effects. However, the delay of one time bin also exists in the full bilingual data set (English and Dutch combined), which is equal in size to the monolingual data set. This supports

[12] An interesting way to assess whether prediction was affected by speaker accent, is to look at the prediction effect throughout the course of the English part of the experiment. If there was an effect of accent, listeners may have adapted to the speaker accent throughout the English block (although other factors such familiarity with the task and experimental design may also yield such adaptations). We checked whether there was an effect of the time course of the experiment by testing whether the effect of condition was larger in the second half of the English block than in the first half of the English block, both for bilinguals and monolinguals. The factor 'experiment half' (first half vs. second half) was added to the model used for the analysis of the English (monolingual and bilingual) data. The interaction effect between experiment half and condition (constraining vs. neutral) was not significant (β = .03, SE = .21, t= -.16, p=.87). The main effect of experiment half was also not significant (β =-.41, SE = .22, t= -1.90, p=.07). No difference was found between the proportion of fixations on the target image or predictive behavior in the first and second half of the English block.

that there may not just a power issue, but that in fact bilinguals predicted upcoming information slightly less rapidly than monolinguals. This would be consistent with the weaker links hypothesis of bilingual language processing, which states that division of use between a bilingual's two languages results in weaker links between lexical items' semantics and phonology (Gollan et al., 2008). This should result in slower lexical access and could possibly lead to slower predictions during language comprehension.

This study shows that L2 listeners use semantic information provided by sentences to restrict the expected subsequent domain of reference to the same extent as in L1 processing by bilinguals and monolinguals. This finding suggests that, when no grammatical rules need to be processed online in order for participants to generate a prediction, the basic principles of recent theories of prediction in language comprehension (cf. Altmann & Mirković, 2009; Federmeier, 2007; Kutas et al., 2011; Pickering & Garrod, 2013) also apply to L2 processing in highly proficient bilinguals. Future studies will have to point out more precisely in what circumstances predictive language processing is retained in L2 processing, and when it is not.

REFERENCES

Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 520.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition, 73,* 247–264.

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and Prediction in Human Sentence Processing. *Cognitive Science, 33,* 583–609.

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,* 390–412

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11,* 280–289.

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59,* 457–474.

Boland, J. E. (2005). Visual arguments. *Cognition, 95,* 237–274.

Broersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer (Version 5.3.78). Retrieved from www.praat.org

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 1029–1040.

Cook, V. (1997). The consequences of bilingualism for cognitive processing. In A. Groot & J. F. Kroll (eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 279–300). Lawrence Erlbaum.

Dell, G. S., & Chang, F. (2013). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369,* 20120394–20120394.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8,* 1117–1121.

Deyne, S., Navarro, D. J., & Storms, G. (2013). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods, 45,* 480–498.

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). WHEN GENDER AND LOOKING GO HAND IN HAND. *Studies in Second Language Acquisition, 35,* 353–387.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44,* 491–505

FitzPatrick, I., & Indefrey, P. (2007). Effects of sentence context in L2 natural speech comprehension. *Proc. of the Cognitive Neuroscience Master of the Radbout University, 2,* 43–56.

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1461–1469.

Foucart, A., Ruiz-Tada, E., & Costa, A. (2015). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition*, 1–7.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58,* 787–814.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research, 30,* 251–266.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research, 29,* 33–56.

Hopp, H. (2015). Semantics and Morphosyntax in Predictive L2 Sentence Processing. *International Review of Applied Linguistics.*

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition, 96,* 23–32.

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience, 31,* 80–93.

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? In J. Rothman & S. Unsworth (eds.), *Linguistic Approaches to Bilingualism* (pp. 257–282).

Kamide, Y. (2008). Anticipatory Processes in Sentence Processing. *Language and Linguistics Compass, 2,* 647–670.

Koehne, J., & Crocker, M. W. (2015). The Interplay of Cross-Situational Word Learning and Sentence-Level Constraints. *Cognitive Science, 39,* 849–889.

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (ed.), *Predictions in the brain: Using our past to generate a future* (pp. 190–207).

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2013a). Interlingual lexical competition in a spoken sentence context: Evidence from the visual world paradigm. *Psychonomic Bulletin & Review, 20,* 963–972.

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2013b). The influence of sentence context and accented speech on lexical access in second-language auditory word recognition. *Bilingualism: Language and Cognition, 16,* 508–517.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods, 44,* 325–343.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106,* 1126–1177.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*.

Mandera, M., Pawel, Keuleers, E., & Brysbaert, M. (in press). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 843–847.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50,* 940–967.

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language, 69,* 574–588.

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics, 53,* 372–380.

McQueen, J. M., & Huettig, F. (2012). Changing only the probability that spoken words will be distorted changes how they are recognized. *The Journal of the Acoustical Society of America, 131,* 509–517.

Misra, M., Guo, T., Bobb, S. C., & Kroll, J. F. (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language, 67,* 224–237.

Mitsugi, S., & Macwhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition, 19,* 19–35.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36,* 329–347.

R Core Team. (2013). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/.

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, & Psychophysics, 77,* 720–730.

Saslow, M. G. (1967). Latency for saccadic eye movement. *JOSA*, *57*(8), 1030–1033.

Severens, E., Lommel, S. V., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica, 119,* 159–187.

Van Berkum, J. J. A. (2010). The brain is a prediction machine that cares about good and bad—any implications for neuropragmatics. *Italian Journal of Linguistics, 22,* 181–208.

Weber, A., Betta, A. M. D., & McQueen, J. M. (2014). Treack or trit: Adaptation to genuine and arbitrary foreign accents by monolingual and bilingual listeners. *Journal of Phonetics, 46,* 34–51.

Weber, A., & Broersma, M. (2012). Spoken word recognition in second language acquisition. In C. Chapelle (ed.), *The Encyclopedia of Applied Linguistics*. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50,* 1–25.

Woumans, E., Ceuleers, E., Van der Linden, L., Szmalec, A., & Duyck, W. (2015). Verbal and nonverbal cognitive control in bilinguals and interpreters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1579–1586.

Zirnstein, M., Van Hell, J. G., & Kroll, J. F. (2015, May). *How Bilingualism and Cognitive Control Impact L2 Reading Comprehension: Evidence from ERPs*. Poster presented at the International Symposium on Bilingualism.

CHAPTER 3
PREDICTION AND INTEGRATION OF SEMANTICS DURING L2 AND L1 LISTENING[1]

Using the visual world paradigm, we tested whether Dutch-English bilinguals predict upcoming semantic information in auditory sentence comprehension to the same extent in their native (L1) and second language (L2). Participants listened to sentences in L1 and L2 while their eye-movements were measured. A display containing a picture of either a target word or a semantic competitor, and three unrelated objects was shown before the onset of the auditory target word in the sentence. There were more fixations on the target and competitor pictures relative to the unrelated pictures in both languages, before hearing the target word could affect fixations. Also, semantically stronger related competitors attracted more fixations. This relatedness effect was stronger, and it started earlier in the L1 than in the L2. These results suggest that bilinguals predict semantics in the L2, but the spread of semantic activation during prediction is slower and weaker than in the L1.

[1] Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2018). Prediction and integration of semantics during L2 and L1 listening. Manuscript submitted for publication.

INTRODUCTION

Smooth and efficient language comprehension involves prediction of upcoming information. Context information affects the language comprehension system before new bottom-up input is encountered, and this may involve pre-activation of linguistic information (see Kuperberg & Jaeger, 2016 for a recent review; but also see Nieuwland et al., 2017 for a multilab failure to replicate pre-activation of phonology). Linguistic predictions are made on the basis of cues from the linguistic (e.g., Altmann & Kamide, 1999; DeLong, Urbach, & Kutas, 2005; Otten, Nieuwland, & Van Berkum, 2007) and non-linguistsic context information (Chambers, Tanenhaus, & Magnuson, 2004; Salverda, Brown, & Tanenhaus, 2011). The content of predictions also varies greatly. Predictions can consist of semantic properties of upcoming words (including object shape) (e.g., Altmann & Kamide, 1999, 2007; Rommers, Meyer, Praamstra, & Huettig, 2013), syntactic information (e.g., Arai & Keller, 2013), and word form information (e.g., DeLong et al., 2005). Predictive language processing is not an all-or-nothing phenomenon but rather something that occurs in a graded manner (Kuperberg & Jaeger, 2016). Several word candidates for prediction are activated in parallel, depending on how likely they are given the context. Here, we tested whether prediction of target word semantics by bilinguals, and spreading semantic activation to competitors with varying degrees of semantic associatedness, is equally strong in both of their languages.

How much or how strongly a person predicts is affected by individual cognitive differences such as cognitive resources, processing speed (Huettig & Janse, 2016) and language experience (Foucart, 2015; Kaan, 2014; Kuperberg & Jaeger, 2016; Peters, Grüter, & Borovsky, 2015; Phillips & Ehrenhofer, 2015). Each of these factors may differ between a bilingual's native language (L1) and second language (L2), and can therefore potentially affect predictive language processing in each language For example, increased lexical competition due to cross-lingual word coactivation affects speed of lexical acces in bilinguals (Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Lagrou, Hartsuiker, & Duyck, 2013), particularly in L2 (Weber & Broersma, 2012). Bilingual language users usually have much less experience using their L2 than their L1. This may result in weaker links between word forms and semantics (Gollan, Montoya, Cera, & Sandoval, 2008; Gollan et al., 2011) and this may in turn again result in slower or weaker retrieval of linguistic representations. Less use may also result in lower quality of linguistic representations and different frequency biases for prediction, becaue a particular continuation for a prior context may have been encountered less often (Kaan, 2014). Experience may determine how reliable the listener considers her prior knowledge about the linguistic context information to be, which could affect the extent to which the listener engages in prediction (Kuperberg & Jaeger, 2016). Finally, L2 processing may tax working memory more than L1 processing (Francis & Gutiérrez, 2012; McDonald, 2006). Therefore, if working memory resources are requiered for predictive processing (e.g., Huettig & Janse, 2016; Pickering & Gambi, 2018), then prediction may be less efficient in L2 than in L1. In sum, less efficient retrieval of representation due to any of the disadvantages associated with L2 processing may hinder the construction of higher-level meaning (such as sentence meaning) used for generating a prediction. In addition, it may hinder retrieval of the to be predicted representation itself. This may lead to slower, weaker, and/or less accurate predictions.

In a recent account of predictive processing Pickering and Gambi (2018) argue that one route to prediction is optional. It uses covert imitation of the input, construct a representation of speaker intention and engages the production system to generate a targeted prediction (see Dell & Chang, 2013; Huettig, 2015; Pickering & Garrod, 2013, for other accounts assuming involvement of production). The authors hypothesize that this ‘prediction-by-production’ route is likely used less or fails more often in non-native language comprehension because it requieres time and cognitive resources (2018). Prediction-by-production proceeds through the same stages as production. Therefore, later stages of prediction (e.g., syntax and word form) may fail more often than earlier stages (eg., semantics). This account also assumes a second route to prediction, based on spreading activation between associated representations. This ‘prediction-by-association’ route is less accurate than prediction-by-production because it is not targeted, but it is relatively automatic. This entails that it should be mostly intact in populations with limited resources, such as L2 comprehenders.

There is indeed some evidence that the later stages of prediction-by-production sometimes fail in L2. Differences between prediction in L1 and L2 comprehension have been found when a language-specific morpho-syntactic or phonotactic rule needs to be applied quickly and accurately in order to pre-activate a target for prediction or when the target for prediction is word-form (Hopp, 2013, 2015; Ito, Pickering, & Corley, 2018; Martin et al., 2013; Mitsugi & Macwhinney, 2015). For example, in Martin et al.’s (2013) ERP study, native speakers of English and late Spanish-English bilinguals read English sentences with a predictable or unpredictable sentence ending (e.g. *Since it is raining, it is better to go out with an umbrella [EXPECTED]/ a raincoat [UNEXPECTED]*). The article preceding the sentence final noun was always congruent with the final noun, but not always congruent with the expected noun. Martin et al. found an N400-effect on the processing of incongruent versus congruent articles for L1 readers, but not for L2 readers. The sentence final noun elicited an N400-effect as well, in both groups, but the effect was larger for L1 than for L2 readers. Thus, the N400 elicited by the article showed that bilinguals reading in the L2 did not anticipate upcoming word forms like native readers did, but the noun-elicited N400 might indicate that target word integration was easier in both languages when the target word was predictable. Alternatively, it may have been an effect of later prediction in L2, but as it was not measured before the target word the two explanation cannot be teased apart.

Ito et al. (2018) studied prediction of word form using a visual world paradigm. Native English and Japanese-English bilinguals listened to constraining sentences in English (*e.g. The tourists expected rain when the sun went behind the …*), and looked at displays containing either a target object (*cloud*; in Japanese: *Kumo*), a phonological competitor for the target object name in English (*clown*), a phonological competitor for the target object name in Japanese (bear; *kuma*), or an unrelated object (*globe*; *tikyuugi*). Native listeners fixated target objects and English competitors more than distractor objects before hearing the target could affect fixations. Non-native listeners only fixated targets more often than distractors (though later than the native listeners), and not English or Japanese phonological competitors, indicating that they predicted target word semantics but not word form.

Hopp (2015) contrasted prediction based on morpho-syntactic cues and lexico-semantic cues. In a visual world paradigm study, Native German listeners and English-German bilinguals looked at picture displays including three possible actors and a control object while they listened to SVO (e.g. *The$_{NOM}$ wolf kills soon the$_{ACC}$ deer*) or OVS (e.g., *The$_{ACC}$ wolf kills soon the$_{NOM}$ hunter*) sentences in German. Anticipatory looks were found to expected patients (*the deer)* before the onset of the second NP in SVO sentences and at expected agents (*the hunter*) in OVS sentences in the native listener group. On the other hand, the English-German bilinguals were more likely to look at patient objects before the onset of the second NP, independently of first NP case marking (nominative or accusative). Thus, even though Hopp found evidence for prediction based on lexical-semantic cues (verb information) in the L2, no prediction based on morpho-syntactic (case marking) information was found in the L2. Participants' knowledge of the German case marking system was not assessed separately, but German proficiency of the bilingual participants did not affect the pattern of results. Similarly, Mitsugi and MacWhinney (2015) found that English-Japanese bilinguals were unable to use case marking information as a cue for prediction in Japanese, even though the bilinguals' had good offline knowledge of the Japanese case marking system.

The earlier semantic stage of prediction-by-production as well as prediction-by-association are expected to be relatively intact in L2 comprehension, due to these requiering relatively little time and resources. Indeed, when no application of a language-specific (morpho-)syntactic rule is required for prediction (Dijkgraaf et al., 2017; Hopp, 2015; Ito et al., 2017), or when the same rule exists in the participants' L1 (Foucart, Martin, Moreno, & Costa, 2014; Foucart, Ruiz-Tada, & Costa, 2015; van Bergen & Flecken, 2017), L2 listeners often do show prediction effects, like in L1. Dijkgraaf et al. (2017), for example, compared prediction between the L1 and the L2 of the same participants using an eye-tracking paradigm based on Altmann and Kamide (1999). Participants listened to simple SVO sentences with either a constraining (e.g., *Mary knits a scarf*) or a neutral verb (e.g., *Mary loses a scarf*). The visual display showed four objects that could all be lost, but only one that could be knitted (a scarf). Dutch-English participants listening to sentences in Dutch or English were more likely to fixate on the target object in the constraining condition than in the neutral condition, before exposure to the auditory target word could influence fixations. The bias in target fixations did not differ between the L1 and L2. Likewise, using a between-subject comparison, Ito et al. (2017) found that bilinguals listening to constraining and neutral sentences in their L2 (English; various L1 languages) showed similar predictive looking behaviour as L1 listeners. Adding a cognitive load during the listening task (remembering 5 words) affected prediction, but in a similar way for L1 and L2 listeners. These findings indicate that at least in some circumstances, L2 listeners predict upcoming semantic information (be it through prediction-by-association or also in a more targeted way through prediction-by-production). However, as Pickering and Gambi also note, spreading activation in semantic prediction in both routes depends on the number and strength of links between representations (Pickering & Gambi, 2018), which is in turn shaped by (linguistic) experience.

Semantic processing in the L2 may be delayed relative to L1 (see Frenck-Mestre, German, & Foucart, 2014; Moreno, Rodríguez-Fornells, & Laine, 2008 for a review), which may be due to the mapping of L2 words onto semantic memory. Specifically, L1 words may be semantically richer than L2 words, as assumed in different theories of bilingual lexicosemantic memory (Finkbeiner, Forster, Nicol, & Nakamura, 2004; Schoonbaert, Duyck, Brysbaert, & Hartsuiker, 2009). Schoonbaert et al. based their model on the distributed feature model (Van Hell & De Groot, 1998) and suggest that L2 words have less semantic features than L1 words. Therefore, two words that share features in the L1 may have no, or fewer, shared features in the L2. Thus, even though bilinguals are able to make semantic predictions based on lexical-semantic information from the sentence context in the L2, perhaps they do not do so as strongly and as quickly as monolinguals do. This should be the case especially when the semantic associations between the sentence content and the predicted information is weaker, or when remote spreading of activation to concepts semantically associated with the predicted concept is tested. The strength of the links between word forms and semantics may also be weaker in L2 than in L1 (Gollan et al., 2008, 2011), which may similarly affect strength and speed of semantic pre-activation.

In line with this hypothesis, Japanese-English bilinguals listening to constraining sentences showed anticipatory eye-movements to a predictable target object later than English native speakers (e.g., *cloud*, when listening to *The tourists expected rain when the sun went behind the . . .*) (Ito et al., 2018). Also, using ERPs, Ito, Martin & Nieuwland (2017) found no evidence of pre-activation of a semantic competitor of the predictable target word in non-native speakers, whereas such an effect was found in native speakers (Ito, Corley, Pickering, Martin, & Nieuwland, 2016). Similarly, Foucart, Moreno, Martin, and Costa (2015) found that value-inconsistent statements as compared to value-consistent statements (e.g., *Nowadays, paedophilia should be prohibited/tolerated across the world*) triggered an N400 response in native speakers but not in non-native speakers. One possible interpretation of this finding is that the valence of a concept is not retrieved from the word as efficiently in the L2 as in the L1, and that therefore, the L2 speakers did not generate predictions based on concept valence.

Peters, Grüter, and Borovsky (2015) showed that highly proficient bilinguals pre-activated target word semantics faster than low proficient bilinguals. For instance, they fixated pictures of a *ship* faster when listening to the sentence *The pirate chases the ship.* In contrast, low-proficient bilinguals were more likely to fixate competitors that were locally related to the action verb, but not necessarily consistent with the sentence meaning (e.g. looking at a *cat* after hearing the verb *chases* in the above sentence. Finally, Kohlstedt and Mani (2018) presented discourse information in a visual world paradigm. When presententing two sentences in which the first contained a semantically associated or a neutral prime for a target in the second, predictive fixations were found in L1 listeners, but not in L2 listeners. However, eventhough the pattern of results differed for each group, the overall difference between groups (bilinguals in L2 vs. native speakers) in the effect of context (biasing or neutral) on target fixations was not significant.

In sum, bilinguals can predict upcoming information during L2 processing in some circumstances, but they do not always do so to a similar extent as native speakers when application of a language specific morpho-syntactic or phonotactic rule is required. In addition, even though some research suggests that lexical-semantic prediction is intact in bilinguals, there is also evidence suggesting that lexical-semantic prediction is affected in bilinguals comprehending L2 input. We hypothesize that even though lexical-semantic prediction can occur in L2 comprehension, the inconsistent findings above may be due to differences in spreading semantic activation and/or temporal dynamics between L1 and L2, with differences especially arising in more challenging contexts. Here, we will investigate when and how prediction in L2 differs from L1, using targets that vary in predictability, and how spreading semantic activation evolves differently when listening in different languages. More specifically, we expect pre-activation of semantic competitors of expected words to be weaker and/or slower in the L2 than in the L1, especially when the semantic distance between expected words and semantic competitors is larger. That is, we expect prediction to be semantically narrower in the L2. If L2 words are indeed mapped onto fewer semantic features than L1 words (Schoonbaert et al., 2009), they also activate fewer features shared with semantically associated concepts, which should trigger less activation spreading to those concepts in L2.

The Present Study

In the present experiment, we used the visual world paradigm to test whether prediction of semantic information during auditory speech recognition, based on lexical-semantic information from the sentence context, is weaker and/or slower in the L2 than in the L1. Dutch-English bilinguals listened to sentences in Dutch and in English while they looked at four-picture displays on a screen in front of them. The picture display included three items that were unrelated to the target word and an experimental image: either a depiction of the target word or of a semantically related competitor. The semantic distance between the target word and the semantic competitor varied. This way, we were able to test in a more refined way whether prediction in the L1 vs. the L2 leads to a different degree of spreading semantic activation. If this were the case, one would expect a different effect of semantic distances between targets and competitors in each language. Ito et al. (2017) also included a semantic competitor in a visual world paradigm experiment in which they compared prediction in the L1 and L2. However, no pre-activation of the semantic competitor was found in either the L1 or the L2. The absence of an effect of pre-activation may have been caused by the fact that the picture displays in that study included both a target object and a semantic competitor, so that the target object attracted looks so strongly that it prevented any looks to the competitor object (Huettig & Altmann, 2005; Huettig, Rommers, & Meyer, 2011). As a more sensitive measure of competitor activation, we therefore opted for a design in which either the target object or the semantic competitor object was present in the display.

Many studies on predictive language processing in the L2 focused on prediction during sentence reading (Foucart et al., 2014; Ito, Martin, et al., 2016; Martin et al., 2013; Molinaro, Giannelli, Caffarra, & Martin, 2017). However, predictive processing may be particularly challenging for non-native speakers in the auditory modality. Speech unfolds over time and therefore a listener cannot go back to the beginning of a sentence like in reading, where the information remains available. Also, misperceptions and misrepresentations of non-native phonemes, a problem that doesn't exist for bilingual reading in the same alphabet, may increase lexical competition during listening comprehension (Weber & Broersma, 2012). Like Dijkgraaf et al. (2017), Foucart et al. (2015), Ito et al. (2017) and Hopp (2015), the current experiment therefore studied predictive processing in the auditory modality.

It is important to note that a comparison of L1 and L2 listening leaves two options: the first is that native listeners are compared with other subjects that listen in the same language, which is however their L2 (e.g. Ito et al., 2017). Even when participant groups are matched on a number of variables such as age, education level and socio-economic status, they may have very different cultural, educational, and linguistic backgrounds. Thus, any differences found between groups may be due to such variables, rather than the experimental factor Language.

The other option is to compare listening in different languages, within the same subjects. Here, we compared listening between L1 and L2 within the exact same Dutch-English bilingual participants. This way, we eliminated confounding effects of individual cognitive differences that may affect prediction such as working memory, processing speed (Huettig & Janse, 2016), age (Federmeier & Kutas, 2005), and verbal fluency (Rommers, Meyer, & Huettig, 2015). This also eliminates the high inter-individual variability that characterizes eye movements (Bargary et al., 2017; Rayner, 1998) and which may confound between-group differences in visual world paradigms. To account for differences between the two languages used in this within-subject design, we included linguistic factors of stimuli such as sentence length, phoneme count, word frequency and semantic distance scores in our analyses.

METHOD

Participants

Bilinguals. 50 native speakers of Dutch took part in the experiment (11 men and 39 women, mean age 19 years, *SD*=2.85). They were Ghent University students participating for course credit.

Dutch was the participants' dominant and most proficient language, and English was their second (49 participants) or third (1 participant) language. On average, participants started acquiring English at age 11 (*SD*=2,46), mainly in school, on holiday or through (online) media. None of the participants had spent time living in an English-dominant country. The participants reported to be exposed to Dutch an average of 73% of the time, and to English 22% of the time. Forty-seven participants also had knowledge of French, and 24 participants had knowledge of German. Nine participants had knowledge of Spanish, two knew Arabic, one Portuguese, and one Italian (all late learners). Language proficiency in English and Dutch was assessed with the LexTALE vocabulary knowledge test (Lemhöfer & Broersma, 2012) and with self-ratings. The LexTALE is a 60-item lexical decision task (unspeeded). It indicates word knowledge and general language proficiency (Lemhöfer & Broersma, 2012). The bilinguals' mean LexTALE scores and self-ratings are reported in Table 1. The participants were significantly less proficient in their L2 than in their L1.

Table 1

Participants' Mean (SD) L1 and L2 LexTALE Scores and Self-ratings

	L1 Dutch	L2 English	p-value[c]
Lextale[a]	88.72 (7.25)	70.05 (10.59)	<0.001
Rating listening[b]	4.98 (.14)	4.00 (.54)	<0.001
Rating speaking[b]	4.94(.32)	3.36 (.60)	<0.001
Rating reading[b]	4.94(.24)	3.78 (.55)	<0.001
Rating general proficiency[b]	4.94 (.24)	3.64 (.55)	<0.001
Category fluency	23.46 (5.23)	14.19 (3.96)	<0.001

[a] Scores consist of percentage correct, corrected for unequal proportion of words and nonwords (Lemhöfer & Broersma, 2012). Due to technical problems one participant's score is missing.

[b] Means of self-assessed ratings on a scale of 1 to 5 (1=not at all, 5=perfect/mother tongue) for listening, speaking, reading and general proficiency.

[c] Reported p-values indicate significance levels of dependent samples t-tests between scores for Dutch and English in bilinguals. Df of t-test on LexTALE scores= 48, Df of t-test on Category Fluency=47 (due to technical problems one participant's LexTALE score and two participants' Fluency scores are missing). Df of all t-tests on ratings= 49.

Materials and Design

Three hundred sixty-two trials were included in the experiment. On each trial, participants listened to a sentence and saw a four-item picture display. Fifty other participants filled out a cloze probability test for an initial set of 871 candidate sentences,[2] with the dual purposes of (a) sentence selection and (b) measuring predictability of sentence-final (target) words. The sentences had varying cloze probabilities (see Figure 1 panel A). Mean cloze probabilities were .71 (*SD*=.23) in Dutch and .68 (*SD*=.24) in English.

The candidate sentences were constructed so that word order was as similar as possible in Dutch and English. Sentences were excluded from the final sentence set if the Dutch and English target provided by the participants were not translation equivalents, and if the provided target word was not depictable or a picture of the word was not included in the normed picture set that we used (Severens, Lommel, Ratinckx, & Hartsuiker, 2005). Also, only one pair of sentences (translation equivalents in Dutch and English) was selected for each target picture. All English sentences were checked for grammaticality by a native speaker of American English. Like the participants in the main experiment, the participants were Ghent University students with knowledge of Dutch (L1) and English (L2). Half of the participants filled out the cloze test for the sentences in Dutch and the other half of the participants filled out the test in English. In the cloze test, participants read each sentence without the sentence-final word and were asked to complete each sentence with the first word that came to mind. For each sentence, the highest cloze probability target was selected in English and in Dutch.

[2] Out of the 871 sentences, 54 were from the Block and Boldwin (2010) sentence set, and 31 from Hamberger, Friedman & Rosen (1996). Another 39 were adapted from Block and Boldwin, and 31 were adapted from Hamberger, Friedman and Rosen. These sentences were adapted so that they could be translated to Dutch without changing the sentence final word.

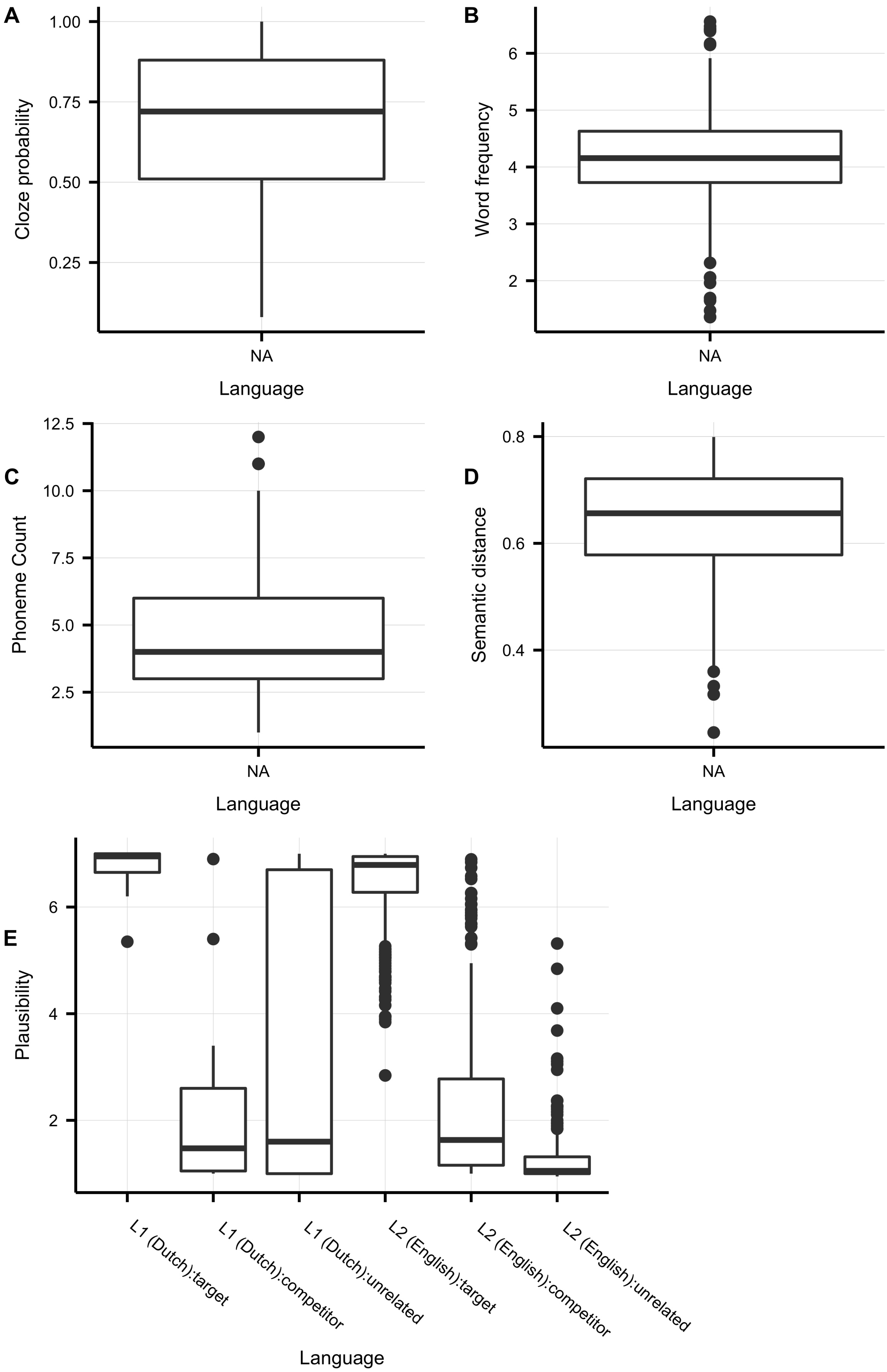

A
1.00
0.75
0.50
0.25
Cloze probability
NA
Language
B
6
5
4
3
2
Word frequency
NA
Language
C
12.5
10.0
7.5
5.0
2.5
Phoneme Count
NA
Language
D
0.8
0.6
0.4
Semantic distance
NA
Language
E
6
4
2
Plausibility
L1 (Dutch):target
L1 (Dutch):competitor
L1 (Dutch):unrelated
L2 (English):target
L2 (English):competitor
L2 (English):unrelated
Language

Figure 1. Stimulus information. A. Stimulus Sentence Cloze Probability. B. Target word frequency. Zipf value (log10(frequency per million*1000)) retrieved from the SUBTLEX-UK and SUBTLEX-NL databases (Keuleers, Brysbaert, & New, 2010; Van Heuven, Mandera, Keuleers, & Brysbaert, 2014). Please note that for six compound nouns no frequency score was available for English. C. Target word phoneme count retrieved from CELEX database (Baayen, Piepenbrock, & Gulikers, 1995). D. Semantic Distance Target-Competitor Pairs Extracted From SNAUT (Mandera, Keuleers, & Brysbaert, 2017). E. Plausibility ratings of target, competitors, and unrelated words as sentence endings. Ratings were given on a 7 point scale ranging from 'not likely at all as sentence ending' to 'very likely as sentence ending'.

Figure 1 panel B and C show the frequency and phoneme count information of the Dutch and English final set of target words. The translation equivalents of the words were mostly phonologically dissimilar in English and Dutch (normalized phonological Levenshtein distance ≤.50, M=.25, SD=.25),[3] but cognates were also included (e.g. L2-L1: *tent-tent, wheel-wiel, nest-nest*), because Dutch and English are related languages and excluding all cognates would lead to unnatural word choices. As phonological similarity between the target word and its translation equivalent may affect looking behaviour, target Levenshtein distance was included as a factor in the analyses and we also confirmed that the data excluding cognates yielded a similar pattern of results.[4] Levenshtein distance between the unrelated picture names and translation equivalents, and between the (auditory) words in the sentences and translation equivalents of each trial may also affect looking behaviour. Given the many English-Dutch cognates and restrictions that had to be taken into account during item construction, we were unable to control for this factor. However, to account for differences in looking behaviour for each item, a random intercept of item was added to the linear mixed models in our analyses.

The pictures in the displays accompanying the sentences were line drawings from the normed database by Severens et al. (2005). Each display accompanying a sentence consisted of either a target picture (the last word in the sentence) or a semantic competitor (a word semantically related to the target word), and three pictures unrelated to the target word. Whether a sentence was

---

[3] 0=no overlap, 1=identical (Schepens, Dijkstra, Grootjen, & van Heuven, 2013).

[4] We applied the optimal models to the prediction time frame data excluding trials in which the experimental image was a cognate (phonological levenshtein distance >.5, following Schepens et al., 2013). For the target, the language by image type interaction remained significant (β = 35, SE = .08, t = 4.19, p < .001). For the competitor data, the threeway interaction between language, image type and semantic distance also remained significant (β = -.21, SE = .08, t = -2.54, p = .01)

accompanied by a target or competitor image was counterbalanced across participants. To ensure that target pictures did not inherently draw more overt visual attention than competitors or unrelated pictures, each of the 362 target pictures was included as a competitor picture for another sentence and as unrelated picture in three other sentences. The 362 experimental sentences thus belonged to 181 sentence pairs. For each sentence pair the target of one sentence was the competitor of the other and vice versa.[5] The display of an experimental trial never included the same picture more than once.

The competitor picture for each target word was selected based on semantic distance scores extracted from the SNAUT database (Mandera et al., 2017).[6] The distance score is based on word co-occurences in large text corpora.[7] The smaller the semantic distance score for a word pair, the more related they are. The score varies between 0 and 1. We included a large range of distance scores for the semantic competitors (see Figure 1 panel D), but the distance score for target-competitor pairs was always smaller than .8. The target-unrelated pairs always had a distance score of more than .8. This cut-off point was chosen because we required a large range of semantic distance scores, and because it was the lowest cut-off point for which it was still possible to pair each target word with the same competitor word in Dutch and in English. Mean semantic distance

[5] The target/competitor words sometimes had false friends in the other language (e.g. *map*, meaning *folder* in Dutch). We applied the optimal models to the prediction time frame data excluding trials in which the experimental image (target or competitor) had (identical) false friends in the other language. Both words with identical orthographic false friends (85 out of 724 words) and words with identical phonological false friends (25 out of 724 words) were excluded (106 in total). For the target, the language by image type interaction remained significant (β = .24, *SE* = .09, t = 2.77, p =.006). As for the competitor, competitor semantic distance still interacted with image type (β = .28, *SE* = .08, t = 3.49, p <.001), but the three-way interaction with language was no longer significant (β =- .13, *SE* = .09, t = -1.54, p =.12). To investigate whether the three-way interaction disappeared because of loss of power or because false friend status actually affected looking behavior we compared the final model with the final model plus the factor false friend status (false friend in the other language yes or no) and the interaction between false friend status and image type. False friend status did not contribute to the model fit ($\chi(2)$=1.73, p=.42).

[6] Competitors were sometimes ungrammatical as sentence ending (e.g. because of a gender mismatch with the preceding determiner) and/or they could violate a phonotactic rule (due to a mismatch with preceding indefinite article *a* or *an*). To test whether competitor grammaticality affected our results we applied the optimal models to the prediction frame data excluding trials in which the competitor was ungrammatical or violated a phonotactic rule. Fifty (out of 362) English sentences and 43 (out of 362) Dutch sentences were excluded. For the target, the language by image type interaction remained significant (β = .25, *SE* = .09, t = 2.89, p =.004). For the competitor data, the twoway language by image type interaction remained significant (β =.22, *SE* = .08, t = 2.68, p =.007), as did the interaction between image type and semantic distance (β =.27, *SE* = .08, t = 3.45, p < .001). The threeway interaction between language, image type and semantic distance approached significance (β = -.15, *SE* = .08, t = -1.87, p = .06). In addition, adding competitor grammaticality and the interaction between grammaticality and image type to the optimal model for the prediction time frame (competitor data set) did not improve the model fit ($\chi(2)$=1.63, p=.44).

[7] The English corpora used were UKWAC (Ferraresi, Zanchetta, Baroni, & Bernardini, 2008) (containing texts from the .uk internet domain) and a subtitle corpus (Mandera, Keuleers, & Brysbaert, 2017) (downloaded from http://opensubtitles.org). For Dutch Sonar-500 text corpus (Oostdijk, Reynaert, Hoste, & van den Heuvel, 2013) (texts from conventional and new media) and another subtitle corpus (Mandera et al., 2017) were used.

score was .63 in Dutch (*SD*=.11) and .64 in English (*SD*=.10). The competitor word never occured in the accompanying sentence.[8] Target and competitor words never started with the same phoneme (except for one pair in Dutch, *orange-lemon*, *sinaasappel-citroen*). As the picture set was limited and each picture had to be used once in every 'position' (target, competitor, unrelated 1, unrelated 2, unrelated 3) it was not possible to take phonetic overlap between unrelated and experimental pictures into account when contructing the picture sets.

Plausibility ratings were generated by 40 further unbalanced Dutch-English bilingual participants (20 in English and 20 in Dutch) for each sentence ending with a target word, a competitor word and with an unrelated word (M=2.14 SD=1.46 on a 7 point scale ranging from 'not likely at all as sentence ending' to 'very likely as sentence ending', see Figure 1, panel E).[9] The participants were recruited from the same Ghent University participant pool, but none of them participated in the cloze probability test nor in the actual experiment. Plausibility was measured after targets were paired with competitors and did not play a role in competitor selection. Competitor plausibility was taken into account in the analyses. Figure 2 shows an example stimulus set, and Appendix 3A contains the sentences and object names of the target and competitor pictures for each stimulus set.

Every twelve experimental sentences were followed by a visually presented simple yes/no question about the preceding sentence to ensure the participants would continue to pay attention to the sentences. To ensure that there were no carry-over effects from answering the question in the data for analysis and to ensure that not every trial would have a target or competitor in the display, we added a filler sentence after each question. The four pictures shown on a filler trial never included a picture of the target word of the accompanying sentence. Unlike the experimental sentences, the filler trials were presented to each participant in Dutch (mean cloze probability=.64) and in English (mean cloze probability=.57). There was no significant difference between the cloze probabilities of the Dutch and English fillers ($t(11)$=1.08, p=.30) The sentences were selected from the same initial candidate sentences as the experimental sentences. The pictures used for the filler trials were not used for the experimental trials.

8 In 8 sentences (out of 362 Dutch and 362 English sentences) either the target word or the competitor word was present in the sentence, either with the same meaning or a slightly different meaning (e.g. *She locked her bicycle to a fence with a lock*, *Ivory is derived from an elephant or a rhino*-> competitor: *elephant*). A picture of the target or competitor word also present in the sentence was likely to attract more fixations in these sentences than in other sentences. The random slope for item in the analyses ensured that this possible confound did not affect the results. In addition, an analysis of the target and competitor data of the full prediction time frame without these 7 sentences did not change the results.

9 Due to an error in the test plausibility ratings for three (out of 724 sentences) were missing.

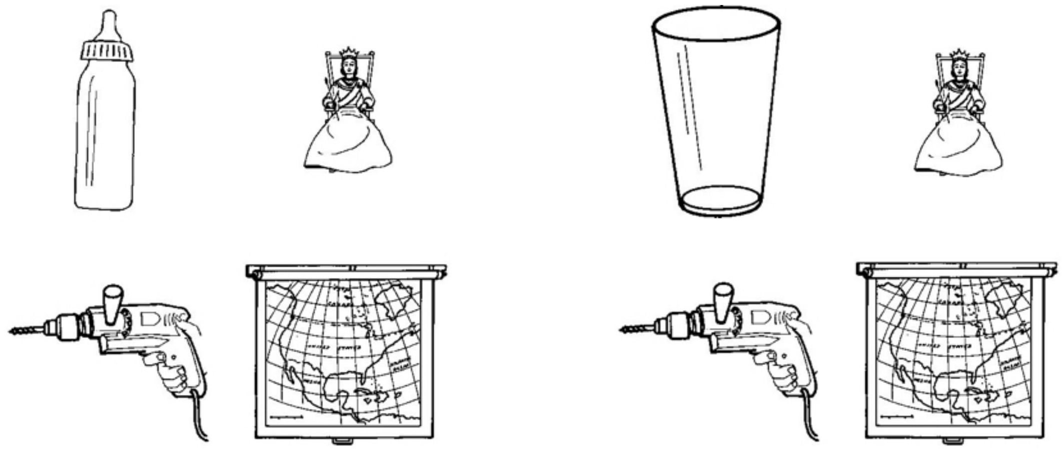


Figure 2. Example stimulus displays. Each participant was presented with one of these two displays with the sentence *'Her baby doesn't like drinking from a bottle'*. The left display includes a picture of the target word for prediction (*bottle*) and the right display includes a picture of a semantic competitor (*glass*). Each display also included 3 unrelated images.

Recordings. The sentences for the experiment were recorded in a sound attenuating room. A Dutch-English bilingual (female, 21 years old) from Flanders who had lived in England from age five to twelve recorded the sentences. The participants in the experiment rated her accent in English as 3.6 and her accent in Dutch as 4.6 on a scale from 1 (very foreign accent) to 5 (native accent). The speaker was asked to pronounce the sentences clearly at a relaxed but natural rate. Each sentence was recorded three times (sampling frequency 48 kHz); the recording that we judged to have the clearest pronunciation and most neutral prosody was selected for the experiment. The average speech rate was 220 words per minute.

The target onset in each sentence was marked using Praat (Broersma & Weenink, 2014). The average target length was 507 ms (range 224-942 ms) in English and 511 (240-1168 ms) in Dutch. On average, the sentence leading up to the target word was 1977 ms in English (range 708-4557 ms) and 2164 ms in Dutch (range 764-4764 ms). Sentence length up to the target was included as factor in the analyses.

Procedure

Participants followed written and oral instructions to listen carefully to Dutch and English sentences and to look at pictures on the screen. They were instructed to look wherever they wanted as long as their gaze did not leave the screen (Huettig & Altmann, 2005; McQueen & Huettig, 2012). In addition, participants were asked to answer the occasional yes/no question about a preceding sentence by pressing "j" for yes and "f" for no. The questions were included to ensure participants continued to listen to the sentences attentively. Participants were presented with 24 questions throughout the experiment (twelve in Dutch and twelve in English). Eye movements were recorded from the right eye with an Eyelink 1000 eye-tracker (SR Research) (1000 Hz) in tower mount.

A fixation cross appeared on the screen for 500 ms, followed by the presentation of a sentence over headphones. Following the procedure in Rommers et al. (2013), the four pictures were presented only 500 ms before the onset of the target word in the sentence. This was done to avoid visual priming of the target or competitor word semantics by the target or competitor picture. Picture location was randomized. After sentence offset, the pictures remained on the screen for 1000 ms. After drift check the next trial started.

The sentence pairs (where one sentence's target was the other sentence's competitor and vice versa) were split into two lists (list A and list B). Each sentence could be presented with a target or a competitor picture and each sentence could be presented in Dutch and in English. The participants were presented with one block of a list with 181 sentences (and 12 fillers) in Dutch and one block of the other list with 181 sentences (and 12 fillers) in English. Language order, list (A or B), and condition (target or competitor) were counterbalanced, resulting in eight presentation lists with a fixed random order. Between the two blocks, eyetracker calibration was repeated. The eye-tracking experiment took approximately one hour.

After the eye-tracking experiment, participants completed the following additional tests: a digit span task, a verbal fluency task, LexTALE Dutch, LexTALE English (Lemhöfer & Broersma, 2012) (see Table 1 for results), and a language background questionnaire based on LEAP-Q (Marian, Blumenfeld, & Kaushanskaya, 2007). The verbal fluency task was performed in Dutch and in English. The participants were asked to name as many words as they could within the categories 'food' and 'animals' within 1 minute. The categories were counterbalanced across languages between participants. Completion of the additional tests took approximately 40 minutes.

Analyses

Our data set was analyzed with linear mixed effects models in R (3.3.2) (R Core Team, 2013) with lme4 (version 1.1-12) (Bates, Mächler, Bolker, & Walker, 2015). The p-values for the fixed effects in our models were obtained using the lmerTest package (version 2.0-33) (Satterthwaite degrees of freedom approximation) (Kuznetsova, Brockhoff, & Christensen, 2016). Post-hoc contrasts were performed with the lsmeans package (Kenward-Roger's approximation to degrees of freedom). Our dependent variable was the empirical logit (a quasi-logit transformation suitable for probabilities that are near 0 or 1) of the proportion of eye-data samples in which there was a fixation to a picture over the total number of samples (Barr, 2008). The proportions of looks to the three unrelated pictures were averaged. We ran separate analyses for the trials in which the display featured the target, and trials in which the display featured a competitor. This was done because the competitor model included the semantic distance factor (semantic distance between the competitor picture name and the target for prediction), whereas the target model did not.

We first analyzed the data of the prediction time frame, without taking into account the time course for prediction. As planning and executing a saccade takes approximately 200 ms (Matin, Shao, & Boff, 1993; Saslow, 1967) the prediction time frame included the eye-data samples starting from 200 ms after the onset of the picture display, to 200 ms after target onset. We also analyzed the data in the time frame starting 200 ms after display onset and ending 1000 ms after target offset (display time frame) to see whether any differences in semantic activation between languages persisted after hearing the target word of the sentence. For these analyses, we first constructed a full model including all theoretically relevant fixed effects and interactions for the prediction time frame (Table 2). The model also included random intercepts of participant and sentence. All continuous predictors were scaled and centered. We then used a backward fitting procedure for the fixed effects (the interaction with the smallest t-value was excluded first), followed by forward fitting the random slopes and then backward fitting fixed effects again to find an optimal model (Barr, Levy, Scheepers, & Tily, 2013). A fixed effect or interaction was excluded if a Chi-square test comparing the model with and without the effect was not significant. We report the results for the optimal model. The optimal models we found for the full prediction time frame for the target and competitor data were then used for a time course analysis, in which we fitted the model for each 50ms time bin in the display time frame (200 ms after display onset up to 1000 ms after target word offset). The data sets and scripts used for the analyses are available online at Open Science Framework (https://osf.io/wy9tm/).

Table 2

Factors and interactions included in the full model for the Target trials and Competitor trials

Fixed factors	Two-way interactions	Three-way interactions
Language (L1 Dutch vs. L2 English)	Language : Image type	Language: Image type: Target onset time
	Language : Target onset time	Language: Image type: Cloze probability
	Language : Cloze probability	Language : Image type : English LexTALE score
	Language : English LexTALE score	
Image type (experimental vs. unrelated)	Image type : Target onset time	
	Image type : Cloze probability	
	Image type : English LexTALE score	
	Image type : experimental image frequency	
	Image type : experimental image phoneme count	
	Image type : experimental image phonetic levenshtein distance	
Target onset time (sentence length upto the target word in ms)		
Cloze probability		
Presentation list		
English LexTALE score		
Experimental image frequency		
Experimental image phoneme count		
Experimental image phonetic levenshtein distance (between L1 and L2 translation equivalents)		
Additional terms competitor model		
Fixed factors	**Two-way interactions**	**Three-way interactions**
Semantic distance (between competitor and target, continuous variable)	Language: Semantic distance	Image type : Language : Semantic distance
	Image type : Semantic distance	
Plausibility (plausibility rating of competitor word as sentence ending)	Image type : Plausibility	

RESULTS

Figure 3 shows the time-course of fixations to target, competitor and unrelated pictures in L1 (Dutch) and L2 (English). The graph shows raw fixation proportions.

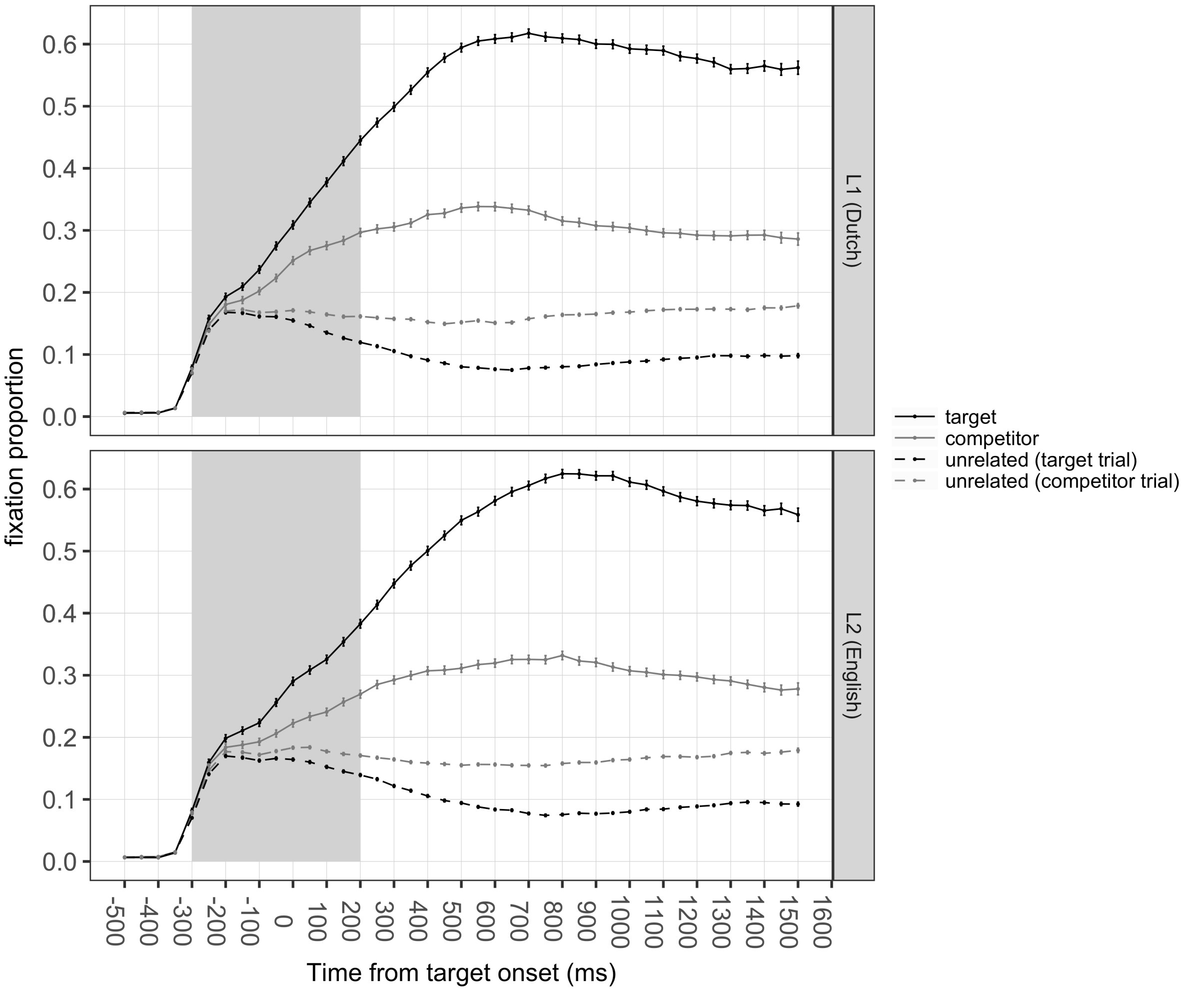


Figure 3. Time course of fixations to target, competitor, and unrelated pictures in the L1 (Dutch) and the L2 (English) relative to target onset. Display onset was 500 ms before target onset. Proportions are based on proportion of samples in which there was a fixation to the picture, aggregated in 50 ms time bins. Proportions for unrelated images were averaged. The area shaded grey is the prediction time frame, in which bottom-up information from the target word could not yet affect looking behaviour (but top-down information from the preceding sentence could). The prediction time frame included the eye-data samples starting from 200 ms after the onset of the picture display to 200 ms after target onset. Whiskers indicate the mean ± standard error.

Visual inspection of the graph suggests that participants were more likely to fixate on target objects than on competitor objects, and also more likely to fixate on competitor objects than on unrelated objects. Fixation proportions for the target, competitor, and unrelated pictures started to diverge well before the target onset time both in Dutch and in English.

Analyses Full Prediction Time Frame

Target trials. The optimal model included the factors language, image type (target versus unrelated), target onset time, and presentation list, as well as the interaction between image type and language, and the interaction between image type and target onset. A random slope of image type was included for each participant and sentence (full results are presented in Table B1 of Appendix 3B).

There was a significant effect of image type (Figure 4, panel A). Importantly, image type also interacted with language. During the prediction time frame, participants were more likely to fixate target images than unrelated images in both the L1 and L2, but in the L1 this effect was larger than in the L2 (β = .26, *SE* = .08, *t* = 3.40, $p < .001$).

The interaction between image type and target onset time was also significant (β = -.38, *SE* = .09, *t* = -4.42, p < .0001). As the length of the sentence leading up to the target word increased, so did the difference between fixations to the target and unrelated images. The interaction between image type and cloze probability did not contribute significantly to the model ($\chi^2(2)$=.28, *p*=.87), suggesting that the bias in looks toward the target picture in the prediction time frame did not increase when the cloze probability of the sentence increased. Also, the interaction between L2 LexTALE score, language, and image type did not contribute significantly to the model ($\chi^2(4)$=4.46, *p*=.35), thus there was no evidence suggesting that relatively proficient bilinguals predicted more than less proficient bilinguals.

Competitor trials. The optimal model included the main effects of language, image type (competitor versus unrelated), semantic distance (between competitor and sentence target, as continuous factor), target onset time, and presentation list. The model also included the two-way interactions between image type and language, image type and target onset, image type and semantic distance, and language and semantic distance. Additionally, the model included the three-way interaction between image type, language, and semantic distance. A random slope of image type was included for each participant and sentence (full results are presented in Table B2 of Appendix 3B).

There was a significant main effect of image type (competitor vs. unrelated) (β = -.66, *SE* = .10, *t* = -6.35, $p < .001$). As shown in Figure 4 panel B, there was a stronger fixation bias to the competitor (versus unrelated images) when the semantic distance between target and competitor was smaller (e.g. *bottle-glass*) (β = .22, *SE* = .07, *t* = 3.04, *p* = .002). This effect was larger in L1 than in L2 (β = -.19, *SE* = .08, *t* = -2.49, *p* = .013). Post-hoc tests reveal that the interaction between semantic distance and image type was significant in both languages (L1 Dutch: $\beta = .66$, $SE = .10$, $t = 6.35$, $p < .0001$, L2 English: $\beta = .51$, $SE = .10$, $t = 4.97$, $p < .0001$).

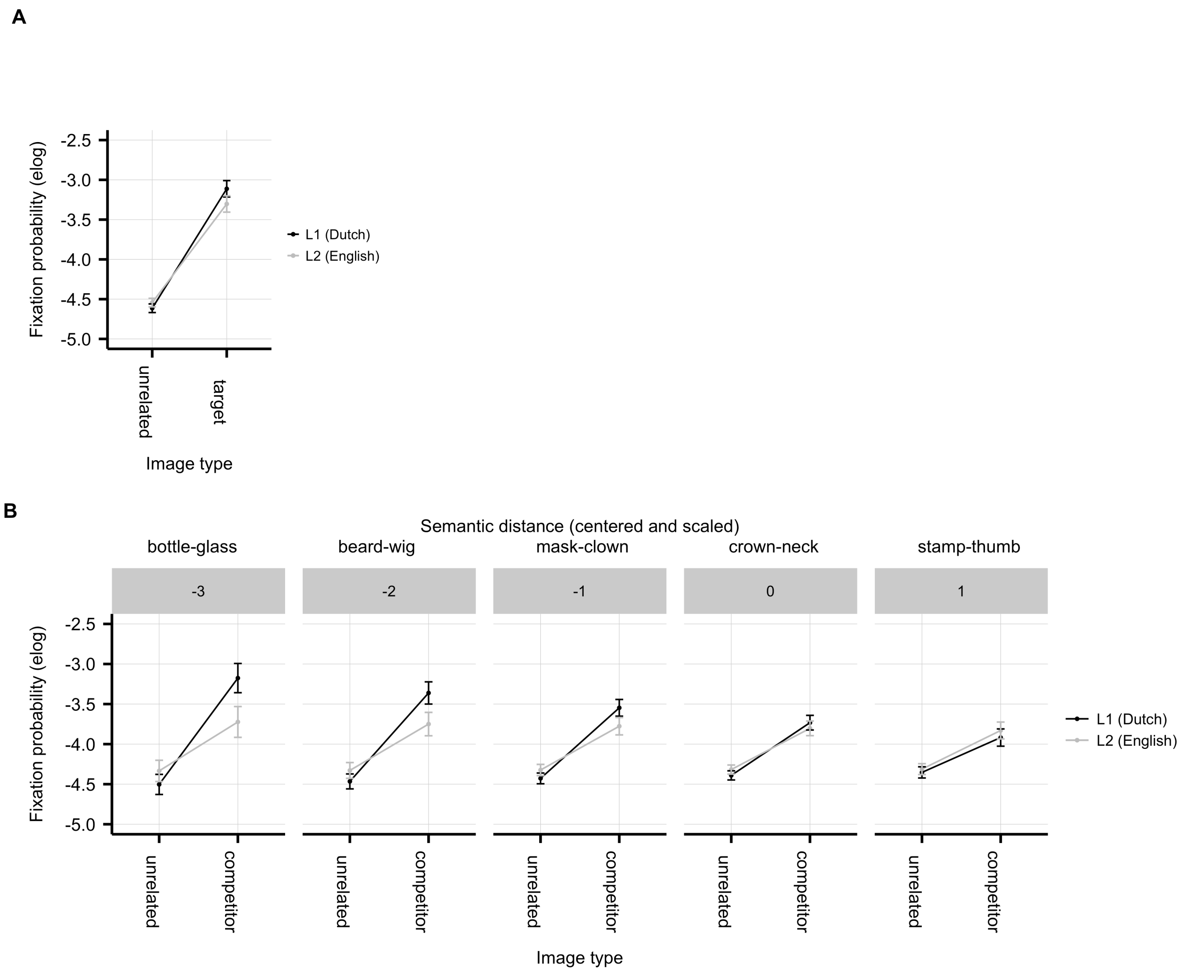


Figure 4. A. Interaction between image type and language for target trials (model predicted means). B. Interaction between image type, language and target-competitor semantic distance (model predicted means). The word pairs above each semantic distance facet are example competitor word pairs in that semantic distance category.

As in the target image data analysis, the interaction between image type and target onset time was significant ($\beta = -.29$, $SE = .08$, $t = -3.57$, $p < .001$). Longer sentences before the target words yielded larger differences between fixations to the competitor and fixations to the unrelated images. As in the target image data, the interaction between image type and cloze probability did not contribute significantly to the model ($\chi^2(2)=1.33$, $p=.51$). Also, the interaction between L2

LexTALE score, language and image type did not contribute significantly to the model ($\chi^2(4)=2.36$, $p=.67$), so that relatively proficient bilinguals did not predict competitors more than less proficient bilinguals.

Individual cognitive differences.

Forward digit span score ($M=9.53$, $SD=1.83$) and fluency (English and Dutch) (Table 1) and their interactions with image type and language did not contribute to the optimal model fit for the competitor and target trials (all *p*s >.1).[10]

Time Course Analyses

A time course analysis was carried out to test whether the language effects found in the analyses of the prediction time frame were caused by a delay in fixation bias in the L2 relative to the L1, rather than by an overall weaker fixation bias in L2. The data were aggregated in 50 ms time bins starting from the prediction time frame (200 ms after the onset of the picture display). The optimal model for the target trials was run for each 50ms time bin in the target trial data, and the optimal model for the competitor trials was run for each 50 ms time bin in the competitor trials. We continued to run the models for the 50 ms time bins after the prediction frame, up to 1500 ms after target onset (the average target duration was 509 ms and pictures were left on screen for 1000 ms after target offset). In those time bins, looking behavior could be influenced by hearing the target. Therefore, we do not interpret the effects in this time window as prediction effects but as effects of ease of integration of information from the auditory target and sentence and the semantic information from the picture display. This type of time-course analysis increases the likelihood of Type I errors, and therefore the differences reported here only include those differences that were found consistently in multiple (>1) time bins (following Ito, Corley, et al., 2017). In addition, we plotted the p-values in each time bin of the most relevant effects with horizontal lines indicating alpha and corrected alpha (Bonferroni style) in Figure C1 and Figure C2 of Appendix 3C.

Figure 5 shows the time course of fixations on the target and unrelated objects in the L1 and L2. The solid circles at the top of the graph indicate a significant interaction between language and image type ($p<.05$).

[10] Due to technical problems the scores for fluency (Dutch and English) is missing for two participants, and the score for digit span is missing for one participant.

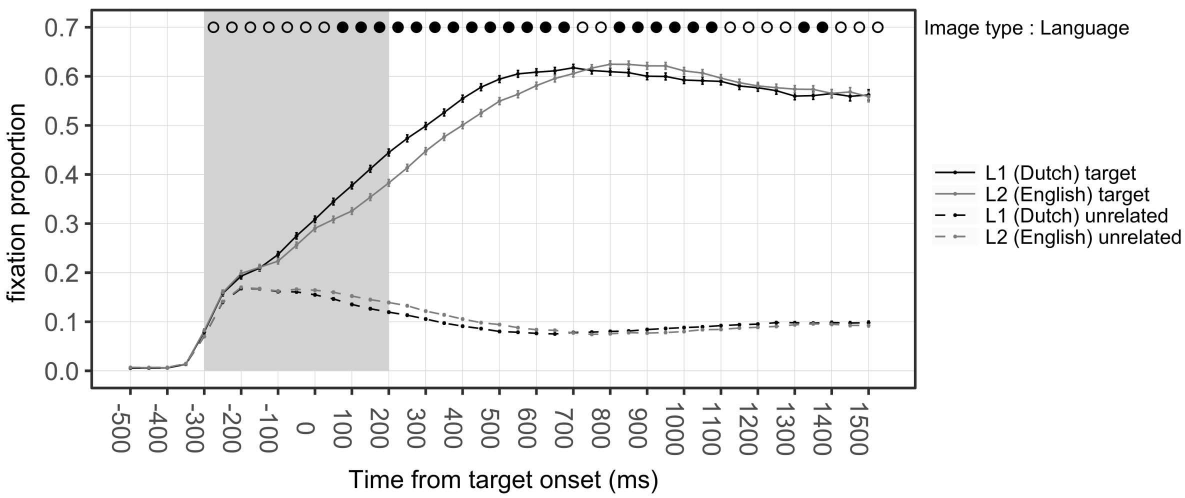


Figure 5. Time course of fixations to the target image and unrelated images in the L1 and L2 relative to target onset. Display onset was 500 ms before target onset. Proportions are based on proportion of samples in which there was a fixation to the picture, aggregated in 50 ms time bins. Proportions for unrelated images were averaged. The area shaded grey is the prediction time frame. Whiskers indicate the mean ± standard error.

In the prediction frame of the target trials, the image type by language interaction was significant only in the last three time bins (50-200 ms after target onset). The main effect of image type (target vs. unrelated) was already significant at 250 ms before target onset. After the prediction time frame, at 700 ms, the bias towards the target did reach the same level in the L2 as in the L1 and from 800 to 1100 ms after target onset the bias towards the target was even larger in the L2 than in the L1.

Figure 6 shows the time course of fixations on the competitor and unrelated objects in the L1 and the L2. The solid circles at the top of the graph indicate a significance of the effects listed on the left (p<.05).

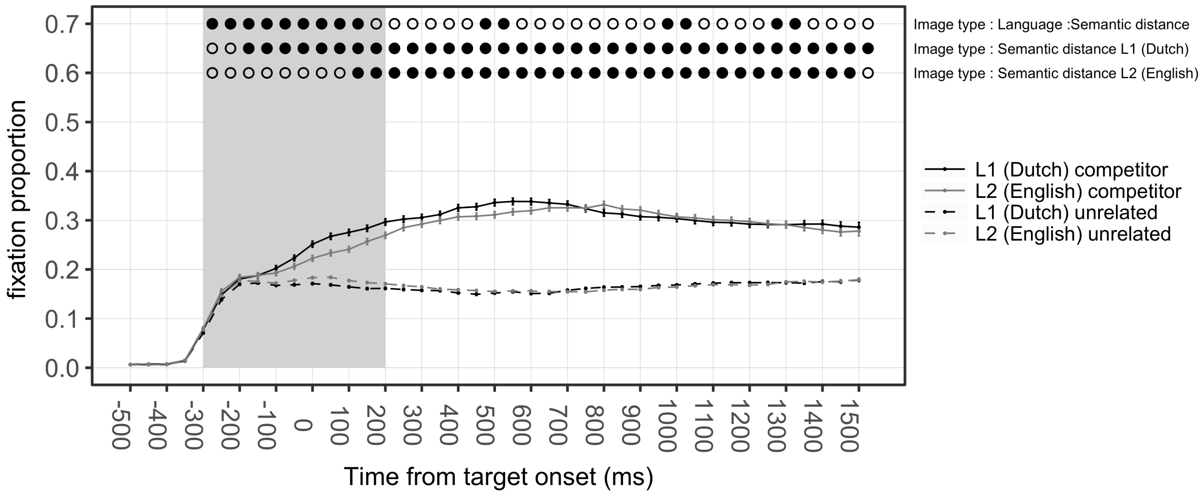


Figure 6. Time course of fixations to the competitor image and unrelated images in the L1 and the L2 relative to target onset. Display onset was 500 ms before target onset. Proportions are based on proportion of samples in which there was a fixation to the picture, aggregated in 50 ms time bins. Proportions for unrelated images were averaged. The area shaded grey is the prediction time frame. Whiskers indicate the mean ± standard error.

First, the main effect of image type became significant 100ms before target onset in the competitor trial data set. The interaction between language and image type was significant from -50 ms to 200 ms in the prediction frame and continued to be significant for 50 ms (200-250 ms) in the post prediction time frame. The bias towards the competitor object was weaker in the L2 than in the L1. The image type effect became significant at 100ms before target on set in both languages separately.

Within the prediction time frame, the interaction between semantic distance and image type was modulated by language from 300 ms before target onset until 150 ms after target onset; the effect of semantic distance on the bias towards the competitor was larger in the L1 than in the L2 in those time bins. Figure D1 of Appendix 3D shows that the interaction effect of semantic distance on the bias towards the competitor gradually increased in the L2 until the three-way interaction with language was no longer significant at 150 ms after target onset. The effect of semantic distance on bias towards the competitor continued to grow in the L2 after the prediction time time frame, and from 450-550 ms, the three-way interaction with language was significant again. This time, the effect of semantic distance on the bias towards the competitor was larger in the L2 than in the L1.

There are four more later time bins in which the the three-way interaction was significant. Again, the effect was larger in the L2 than in the L1 in those time bins. Interestingly, post-hoc tests with lsmeans show that the interaction between image type and semantic distance became significant 300 ms later in the L2 (English) data than in the L1 (Dutch) data (see Figure 6).

Overall Time Course Analysis. In order to compare the time-course for target and competitor pre-activation in both languages we ran an additional time bin analysis on the entire data set, including both target and competitor trials, for the bins in the prediciton time frame. All factors included in both the competitor final model and the target final model were included in the model for the overall analysis. The factor trial type (target vs. competitor) was added as well. Semantic distance was not included as factor as it applied only to the competitor trials. A random slope for image type was added by items and by participants. Further random slopes did not contribute to the model fit (as determined by model comparisons with and without each slope for the model applied to the full prediction time frame data set). The image type effect was significant from 250 ms before target word onset (*p*s<.05), and this effect was modulated by trial type from 150 ms before target onset (*p*s<.05). The bias towards the experimental image was larger on target trials than on competitor trials. The image type effect interacted with language from time bin 0 onwards, with a larger bias towards the experimental image in L1 than in L2. The three-way interaction between image type, trial type and language did not reach significance until the final bin of the prediction time frame. Post-hoc tests reveal that on target trials the effect of image type became significant from 250 ms before target onset onwards in L2, and from 200 ms before target onset in L1. On competitor trials, the effect of image type was significant from 100 ms before target onset onwards in both languages.

DISCUSSION

In the present study, we tested whether prediction of meaning during speech comprehension is affected by language (native versus non-native). We found that bilinguals predicted semantics of a target words both in the L1 and the L2; participants were more likely to focus on target objects than on unrelated objects before the auditory target could affect eye-movements. We found a larger prediction effect when bilinguals listened in the L1 than when they listened in the L2. Bilinguals were also more likely to look at semantic competitor objects than at unrelated objects, in both languages. This shows that semantic pre-activation during listening in both languages is strong enough to spread to related concepts, at least when a picture of the related concept is present on the screen. The strength of the competitor fixation bias depended on the semantic distance between target and competitor (the smaller the distance, the larger the bias) and language: the effect of semantic distance on bias to competitor objects was larger in the L1 than in the L2, with an especially strong competitor effect in the L1 for the most strongly related competitors. Time-course analyses showed that there was significant prediction of target word semantics in the L1 and the L2 250 ms before target word onset, and that the prediction effect was larger in the L1 than in the L2 from 150 ms before auditory exposure to the target word could influence looking behavior. The difference remained significant for 500 ms afterwards. The effect of semantic distance on the bias to competitor objects was larger in the L1 than in the L2 throughout almost the entire prediction time frame. After the prediction time frame, the effect of semantic distance on the bias to the competitor object was the same in the L1 and the L2, and it even became bigger in the L2 than in the L1 for a brief period (6 time bins in total).

In this study, differences were found when directly comparing prediction between the L1 and the L2 of the same individuals when both the cues and information to be predicted are of a lexical-semantic nature. The results indicate that semantic prediction in the L2 does not always occur as efficiently as in the L1. Target pre-activation became significant at approximately the same time in English and Dutch (even one bin earlier in English). This suggests that predictive pre-activation of the target was weaker, rather than slower in L2 than in L2.

The finding that the target object was pre-activated less strongly in the L2 than in L1 differs with earlier findings on semantic prediction in the L2 (Dijkgraaf et al., 2017; Hopp, 2015; Ito et al., 2017). Dijkgraaf et al. directly compared predictive looking behaviour in the L1 and the L2 in bilinguals and found no significant difference. Hopp found predictive looking behaviour in L2 like in L1, but only when the cues used for prediction were lexico-semantic and not when predictions were to be based on case-marking information. No direct comparison of prediction in the L1 and L2 was reported for lexico-semantic prediction. Ito et al. found predictive looking behaviour in the L1 and the L2 but did not directly compare the strength of the prediction effect in each language. Instead, they reported a similar effect of cognitive load on predictive processing in the L1 and L2. Ito et al. (2018) did find an L2 disadvantage in semantic prediction, like we did in the current study. The authors used longer, more naturalistic sentences (e.g., *The tourists expected rain when the sun went behind the cloud*). Both English native speakers and Japanese-English bilinguals showed anticipatory eye-movements to predictable targets (e.g. *cloud*), but the L2-listeners did so later than the L1-listeners.

Prediction during language comprehension is a flexible process that can be modulated by many factors such as the task at hand and individual differences in language experience. Most likely, the differences between our findings and the findings of Dijkgraaf et al. (2017), Ito et al. (2017) and Hopp (2015) can be attributed to contextual factors. The sentences used in the current experiment were longer and often syntactically more complex (e.g. compound sentences) than the simple sentences used in previous studies. This likely hinders predictive processing, in contrast to the above studies where prediction was so straightforward and strong that it occured to the same extent even in a less proficient L2. There may also be a difference in the routes used for prediction in the different studies. Specifically, as in Dijkgraaf et al., Ito et al. (2017), and Hopp predictions were based mainly on information from only one word (the verb), low-level lexical associations may have played a large role. The present study and Ito et al. (2018) used longer, more naturalistic sentences and therefore predictions were likely at least partly based on higher level meaning. If we interpret the findings in the framework of Pickering and Gambi (2018), prediction may have come about mainly via prediction-by-production in the current study and in Ito et al. (2018) , and via prediction-by-association in Ito et al. (2017), and Hopp. Prediction-by-production may be more complex and it may require more cognitive resources unailable to the L2-comprehenders than prediction-via-associations, hence the diverging findings.

Further, in Dijkgraaf et al. (2017), Hopp (2015), and Ito et al. (2017) the picture display appeared before sentence onset. Pre-activation of target word semantics may have been increased greatly because of the visual presence of a plausible target object. This may be especially so for bilinguals, as they may rely strongly on visual information during language processing (Navarra & Soto-Faraco, 2007). Therefore, in order to maximize sensitivity for language differences in the current experiment, the pictures appeared only 500 ms before the onset of the target word in the current experiment.

Besides task and stimulus differences, individual differences between our participants and the participants in the other experiments may also have caused the diverging results. Prediction in the L2 is thought to approach prediction in the L1 as L2 proficiency increases (Kaan, 2014). However, participants in Ito et al. (2017), and Dijkgraaf et al. (2017) were highly proficient like the participants in the current experiment, which makes proficiency an unlikely explanation for the diverging results. Also, like in Ito et al., Hopp (2015) and Dijkgraaf et al., no effect of proficiency on semantic prediction in L2 was found in the current experiment. Perhaps the range of proficiencies was too small to detect such an effect.

Finally, the present experiment had 362 stimulus sentences versus 16 sentence pairs in Ito et al. (2017) and 16 sentences in Hopp (2015). Fewer participants were tested in our study than in the other two, but our comparison across languages was within-participants. Thus, we may have found an effect of language (L1 vs. L2) here because we had more statistical power due to our design and very large number of stimuli.

Our finding that the semantic distance effect on competitor prediction was smaller in the L2 than in the L1 in the prediction time frame indicates that spread of semantic activation due to target pre-activation started later in the L2 than in the L1, that activation spreading was weaker (especially for the most strongly related concepts), or both.

The first explanation receives support from the time-course analyses of competitor trials, which indicated that the effect of spread of semantic activation became significant later in the L2 than in the L1. When we compared looking behavior in the L1 and L2 in later time bins (including time bins where hearing the target word could affect looking behaviour) the effect of semantic distance on the bias to the competitor was the same in both languages, or even bigger in the L2. The later significant effect in the L2 suggests a delay in activation. This would be consistent with the temporal delay assumption of the BIA+ model of bilingual visual word recogntion (Dijkstra & van Heuven, 2002). This assumption states that due to lower subjective L2 word frequency, activation of word form and, as a consequence, semantic codes is somewhat delayed in the L2 compared to the L1, while activation patterns themselves are the same.

We also obtained evidence supporting the second explanation above, namely that of weaker lexico-semantic activation in the L2. We observed that the semantic distance effect in the competitor trials was stronger in the L1 than the L2. We predicted such an effect from the assumption that L2 words are mapped onto fewer semantic features than L1 words (Schoonbaert et al., 2009; Van Hell & De Groot, 1998), and that therefore spreading semantic activation should be narrower in the L2 than in the L1. We expected that the diffence between the L1 and L2 would be particularly large for less strongly related competitors, because L2 concepts should map onto the core semantic features (shared by strongly related concepts), but perhaps not onto the more remote ones (shared by weakly related concepts). Somewhat surprisingly, the difference between the competitor effects in L1 and L2 was most pronounced for the most strongly related competitors, with very strong semantic pre-activation of closely related concepts especially from L1 words. This suggest that stronger spreading semantic activation for the L1 is determined by the strength of mappings between word forms and semantics, rather than by the number of mapped semantic features. Our interaction effect between language, image type and semantic distance suggests that L1 words have stronger links with the underlying concepts than L2 words, which then leads to stronger semantic pre-activation for very related concepts. Such an explanation is consistent with for instance the weaker links account, which assumes that divided language practice across languages leads to weaker links between representations in the bilingual language system (Gollan et al., 2008; Gollan, Montoya, Fennema-Notestine, & Morris, 2005). Because L2 exposure is far less frequent for our bilinguals, mappings from L2 word forms onto semantics are weaker.

As less cognitive resources may be available during L2 than during L1 processing (e.g., Francis & Gutiérrez, 2012; McDonald, 2006) we expected participants with a larger working memory capacity to have less of a disadvantage in L2 prediction. In contrast, we found no effects of working memory span (forward digit span) and verbal fluency score on prediction in L1 and L2, suggesting that working memory resources may not be drive the differences between L1 and L2. Consistent with our finding, Ito et al. found that a cognitive load during speech comprehension affects prediction in L1 and L2 to the same extent. However, the sample of 50 participants in this study may not have been large enough to detect an effect of individual differences in working memory capacity. Future research using a more sensitive design could be aimed at testing whether working memory resource limitations in L2 may underlie the L2 disadvantage in prediction.

For both the target and the competitor data we found that target onset time (the length of the sentence leading up to the target) affected prediction. The longer the sentence, the larger the prediction effect. This may be due both to the increased time for pre-activation in longer sentences and the increased amount of context information to serve as cue for prediction. The effect of sentence length on predictive looking behavior was not modulated by language (L1 vs. L2). Apparently, even though semantic pre-activation was weaker in the L2 than in the L1, the length of the sentence did not differentially affect pre-activation in the L1 and the L2. A limitation of the current study is that the Dutch sentences were slightly longer than the English sentences, possibly contributing to the L2 disadvantage in prediction. However, note that we found an effect of language in addition to an effect of length.

Somewhat unexpectedly, we found no effect of sentence cloze probability on target or competitor pre-activation, even though we included sentences with a rather large range of cloze probabilities (0.08-1). The cloze probability test was filled out with the sentences as context only.

The presence of a picture display with a target or competitor word may have increased the probability of the sentence ending with the target word, thereby eliminating the cloze probability effect. Furthermore, participants listened to 362 experimental sentences with an average cloze probability of .68 for English and .71 for Dutch. The exposure to so many predictable sentences may have further enhanced the likelihood of predictive behavior overall (Lau, Holcomb, & Kuperberg, 2013), and thereby reduced the chances of finding an effect of cloze probability. Finally, superficial lexical associations, rather than the full sentence meaning may have contributed to pre-activation of target and competitor word semantics (Chow, Lau, Wang, & Phillips, in press; Chow, Smith, Lau, & Phillips, 2016; Phillips & Ehrenhofer, 2015). Cloze probability may be affected by both lexical associations and full sentence meaning, but it should be a more exact measure of the latter.

Finally, in this paradigm, we cannot distinguish between competitor activation through target word pre-activation, followed by spreading activation to the competitor on the one hand, and competitor activation via passive resonance of the semantics of semantically related words in the sentence on the other hand. Both mechanisms may also be additive. Future studies could be aimed at pinpointing the exact locus of the delay in/weaker effect of spreading semantic activation in L2 compared to L1. In any case, the present results show that L2 yields slower and/or weaker semantic prediction overall.

In sum, even in an experimental setting with many relatively high cloze sentences and additional visual information, we find differences in the strength and time-course between L1 and L2 semantic prediction. Therefore, language dominance (L1 versus L2) can not only affect prediction based on (morpho-)syntactic cues but also prediction of semantic information based on semantic context information. The difference between prediction in the L1 and the L2 is compatible with the hypothesis that lexico-semantic mappings are weaker for L2 than for L1 (Gollan et al., 2008, 2005), and with slower word form activation and, as as a consequence, slower spread of semantic activation in L2 than in L1, due smaller subjective word frequency in the L2 (Dijkstra & van Heuven, 2002). As working memory (digit span score) did not affect prediction, an explanation in terms of limited cognitive resources in L2 (Francis & Gutiérrez, 2012; McDonald, 2006) is less likely. We suggest that there is no qualitative difference between lexico-semantic prediction in the L2 and the L1, but that subtle quantitative differences arise when graded semantic relations are assessed, like in the present paradigm. The differences between our findings and previous research in which no language effect on semantic prediction was found, illustrate again that prediction during language comprehension is a highly flexible process. Future studies should be aimed at testing which exact contextual factors and individual differences, best explain the diverging findings on predictive behavior in L2 comprehension.

REFERENCES

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*(4), 502–518. https://doi.org/10.1016/j.jml.2006.12.004

Arai, M., & Keller, F. (2013). The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes*, *28*(4), 525–560. https://doi.org/10.1080/01690965.2012.658072

Baayen, H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database. Release 2 (CD-ROM)*. Linguistic Data Consortium, University of Pennsylvania.

Bargary, G., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). Individual differences in human eye movements: An oculomotor signature? *Vision Research*, *141*, 157–169. https://doi.org/10.1016/j.visres.2017.03.001

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*(4), 457–474. https://doi.org/10.1016/j.jml.2007.09.002

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Block, C. K., & Baldwin, C. L. (2010). Cloze probability and completion norms for 498 sentences: Behavioral and neural validation using event-related potentials. *Behavior Research Methods*, *42*(3), 665–670. https://doi.org/10.3758/BRM.42.3.665

Broersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer (Version 5.3.78). Retrieved from www.praat.org

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and Affordances in Syntactic Ambiguity Resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(3), 687–696. https://doi.org/10.1037/0278-7393.30.3.687

Chow, W.-Y., Lau, E., Wang, S., & Phillips, C. (in press). Wait a second! Delayed impact of argument roles on on-line verb prediction. *Language, Cognition, and Neuroscience*.

Chow, W.-Y., Smith, C., Lau, E., & Phillips, C. (2016). A "bag-of-arguments" mechanism for initial verb predictions. *Language, Cognition and Neuroscience*, *31*(5), 577–596. https://doi.org/10.1080/23273798.2015.1066832

Dell, G. S., & Chang, F. (2013). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120394–20120394. https://doi.org/10.1098/rstb.2012.0394

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121. https://doi.org/10.1038/nn1504

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition*, *20*(5), 917–930. https://doi.org/10.1017/S1366728916000547

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(03). https://doi.org/10.1017/S1366728902003012

Duyck, W., Assche, E. V., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 663–679. https://doi.org/10.1037/0278-7393.33.4.663

Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, *42*(2), 133–141. https://doi.org/10.1111/j.1469-8986.2005.00274.x

Ferraresi, A., Zanchetta, E., Baroni, M. G., & Bernardini, S. (2008). Introducing and evaluating ukWaC, a very large web-derived corpus of English. In *Proceedings of the 4th Web as Corpus Workshop (WAC-4)* (pp. 47–54).

Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language*, *51*(1), 1–22. https://doi.org/10.1016/j.jml.2004.01.004

Foucart, A. (2015). Prediction is a question of experience. *Linguistic Approaches to Bilingualism*, *5*(4), 465–469. https://doi.org/10.1075/lab.5.4.04fou

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1461–1469. https://doi.org/10.1037/a0036756

Foucart, A., Moreno, E., Martin, C. D., & Costa, A. (2015). Integration of moral values during L2 sentence processing. *Acta Psychologica*, *162*, 1–12. https://doi.org/10.1016/j.actpsy.2015.09.009

Foucart, A., Ruiz-Tada, E., & Costa, A. (2015). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition*, 1–7. https://doi.org/10.1017/S1366728915000486

Francis, W. S., & Gutiérrez, M. (2012). Bilingual recognition memory: Stronger performance but weaker levels-of-processing effects in the less fluent language. *Memory & Cognition*, *40*(3), 496–503. https://doi.org/10.3758/s13421-011-0163-3

Frenck-Mestre, C., German, E. S., & Foucart, A. (2014). Qualitative Differences In Native And Nonnative Semantic Processing as Revealed by ERPs. In *Foundations of Bilingual Memory* (pp. 237–255). Springer, New York, NY. https://doi.org/10.1007/978-1-4614-9218-4_12

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, *33*(7), 1220–1234. https://doi.org/10.3758/BF03193224

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209. https://doi.org/10.1037/a0022256

Hamberger, M. J., Friedman, D., & Rosen, J. (1996). Completion norms collected from younger and older adults for 198 sentence contexts. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 102–108. https://doi.org/10.3758/BF03203644

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*(1), 33–56. https://doi.org/10.1177/0267658312461803

Hopp, H. (2015). Semantics and Morphosyntax in Predictive L2 Sentence Processing. *International Review of Applied Linguistics.* https://doi.org/10.1515/iral-2015-0014

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, *96*(1), 23–32. https://doi.org/10.1016/j.cognition.2004.10.003

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*(2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003

Ito, A., Corley, M., & Pickering, M. J. (2017). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, *21*(2), 251–264. https://doi.org/10.1017/S1366728917000050

Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language*, *86*, 157–171. https://doi.org/10.1016/j.jml.2015.10.007

Ito, A., Martin, A. E., & Nieuwland, M. S. (2017). On predicting form and meaning in a second language. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*, *43*(4), 635–652. https://doi.org/10.1037/xlm0000315

Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, *98*, 1–11. https://doi.org/10.1016/j.jml.2017.09.002

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? In *Linguistic Approaches to Bilingualism* (pp. 257–282).

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643–650. https://doi.org/10.3758/BRM.42.3.643

Kohlstedt, T., & Mani, N. (2018). The influence of increasing discourse context on L1 and L2 spoken language processing. *Bilingualism: Language and Cognition*, *21*(01), 121–136. https://doi.org/10.1017/S1366728916001139

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: Tests in Linear Mixed Effects Models*. Retrieved from https://CRAN.R-project.org/package=lmerTest

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2013). Interlingual lexical competition in a spoken sentence context: Evidence from the visual world paradigm. *Psychonomic Bulletin & Review*, *20*(5), 963–972. https://doi.org/10.3758/s13423-013-0405-4

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience*, *25*(3), 484–502. https://doi.org/10.1162/jocn_a_00328

Lemhöfer, K., & Broersma, M. E. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, *92*, 57–78. https://doi.org/10.1016/j.jml.2016.04.001

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, *53*(4), 372–380. https://doi.org/10.3758/BF03206780

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*(3), 381–401. https://doi.org/10.1016/j.jml.2006.06.006

McQueen, J. M., & Huettig, F. (2012). Changing only the probability that spoken words will be distorted changes how they are recognizeda). *The Journal of the Acoustical Society of America*, *131*(1), 509–517. https://doi.org/10.1121/1.3664087

Mitsugi, S., & Macwhinney, B. (2015). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, *19*(01), 19–35. https://doi.org/10.1017/S1366728914000881

Moreno, E. M., Rodríguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics*, *21*(6), 477–508. https://doi.org/10.1016/j.jneuroling.2008.01.003

Navarra, J., & Soto-Faraco, S. (2007). Hearing lips in a second language: visual articulatory information enables the perception of second language sounds. *Psychological Research*, *71*(1), 4–12. https://doi.org/10.1007/s00426-005-0031-5

Nieuwland, M., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., … others. (2017). Limits on prediction in language comprehension: A multi-lab failure to replicate evidence for probabilistic pre-activation of phonology. *BioRxiv*, 111807.

Oostdijk, N., Reynaert, M., Hoste, V., & van den Heuvel, H. (2013). *SoNaR user documentation*. Retrieved from https://ticclops.uvt.nl/SoNaR_end-user_documentation_v.1.0.4.pdf

Otten, M., Nieuwland, M. S., & Van Berkum, J. J. (2007). Great expectations: Specific lexical anticipation influences the processing of spoken language. *BMC Neuroscience*, *8*(1), 89. https://doi.org/10.1186/1471-2202-8-89

Peters, R., Grüter, T., & Borovsky, A. (2015). Anticipatory and Locally Coherent Lexical Activation Varies as a Function of Language Proficiency. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 1865–1870). Austin, TX.

Phillips, C., & Ehrenhofer, L. (2015). The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism*, *5*(4), 409–453. https://doi.org/10.1075/lab.5.4.01phi

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: a theory and review. *Psychological Bulletin*. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(04), 329–347. https://doi.org/10.1017/S0140525X12001495

R Core Team. (2013). *R: A language and environment for statistical computing.* Vienna, Austria.: R Foundation for Statistical Computing,. Retrieved from http://www.R-project.org/.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. http://dx.doi.org/10.1037/0033-2909.124.3.372

Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia*, *51*(3), 437–447. https://doi.org/10.1016/j.neuropsychologia.2012.12.002

Salverda, A. P., Brown, M., & Tanenhaus, M. K. (2011). A goal-based perspective on eye movements in visual world studies. *Acta Psychologica*, *137*(2), 172–180. https://doi.org/10.1016/j.actpsy.2010.09.010

Saslow, M. G. (1967). Latency for Saccadic Eye Movement*. *JOSA*, *57*(8), 1030–1033. https://doi.org/10.1364/JOSA.57.001030

Schepens, J., Dijkstra, T., Grootjen, F., & van Heuven, W. J. B. (2013). Cross-Language Distributions of High Frequency and Phonetically Similar Cognates. *PLoS ONE*, *8*(5), e63006. https://doi.org/10.1371/journal.pone.0063006

Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2009). Semantic and translation priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition*, *37*(5), 569–586. https://doi.org/10.3758/MC.37.5.569

Severens, E., Lommel, S. V., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, *119*(2), 159–187. https://doi.org/10.1016/j.actpsy.2005.01.002

van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language*, *92*, 26–42. https://doi.org/10.1016/j.jml.2016.05.003

Van Hell, J. G., & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, *1*(3), 193–211. https://doi.org/10.1017/S1366728998000352

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: a new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology (2006)*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Weber, A., & Broersma, M. E. (2012). Spoken word recognition in second language acquisition. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Hoboken, NJ, USA: John Wiley & Sons, Inc. Retrieved from http://doi.wiley.com/10.1002/9781405198431

CHAPTER 4
PREDICTION OF SEMANTICS IN NATIVE AND NON-NATIVE SPEECH COMPREHENSION: THE ROLE OF COGNITIVE LOAD AND PROCESSING SPEED[1]

The goal of this study was to test whether cognitive load or processing speed explain L2 disadvantages in prediction. Dutch-English bilinguals listened to sentences in English and Dutch and looked at a display presented shortly before the sentence-final target word, while their eye-movements were measured. The display contained a picture of the target object or a semantic competitor, and three unrelated objects. Cognitive load (Experiment 1) and speech rate (Experiment 2) were experimentally manipulated. An additional cognitive load reduced predictive eye-movements to targets (and not competitors) in both languages, but the load effect was larger in L1. Faster L1 speech led to weaker target (but not competitor) prediction compared to normal L1 speech, and competitor (but not target) prediction in L2 was enhanced by slower rate. The results are consistent with the view that bilinguals rely less on resource intensive routes to prediction in L2 than in L1.

[1] Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2018). Modulating prediction of semantics in native and non-native speech comprehension: the role of cognitive load and processing speed. Manuscript submitted for publication.

INTRODUCTION

Native (monolingual) speech processing is usually fast and efficient. One mechanism that supports such smooth language comprehension is the prediction of upcoming information (e.g., Kuperberg & Jaeger, 2016; Van Berkum, 2010). A growing body of evidence suggests that like native comprehenders, non-native comprehenders can predict upcoming semantic (Chambers & Cooke, 2009; Dijkgraaf, Hartsuiker, & Duyck, 2017; Hopp, 2015; Ito, Corley, & Pickering, 2017), and syntactic (Foucart, Martin, Moreno, & Costa, 2014; Foucart, Ruiz-Tada, & Costa, 2015) information. However, many studies have also found weaker, slower or non-significant prediction in the L2 (Dijkgraaf, Hartsuiker, & Duyck, 2018; Hopp, 2013, 2015; Ito, Pickering, & Corley, 2018; Kaan, Kirkham, & Wijnen, 2014; Martin et al., 2013; Mitsugi & Macwhinney, 2015; Sagarra & Casillas, 2018; van Bergen & Flecken, 2017). Here, we tested whether the difference between prediction in L1 and L2 found in some studies can be explained by either lower availability of cognitive resources or slower processing speed.

Differences between predictive processing in L1 and L2 especially seem to be found when (morpho-)syntactic information is used as a cue for prediction. For example, Hopp (2015) used the visual world paradigm to test whether bilinguals could use L1-specific case-marking information to predict upcoming referents in L2. Native German participants and English-German unbalanced bilinguals listened to German sentences with case-marked articles such as *The$_{NOM}$ wolf kills soon the$_{ACC}$ deer* (SVO) or *The$_{ACC}$ wolf kills soon the$_{NOM}$ hunter* while they looked at displays depicting three possible actors and a control object. Native listeners launched anticipatory looks to likely patients in SVO sentences (the deer), and to likely agents (the hunter) in OVS sentences. The non-native participants were more likely to fixate patients before the onset of the second noun phrase in the sentence both in SVO and OVS sentences, indicating that they used semantic information to predict likely upcoming referents, but that they did not use case-marking information to adjust their expectations of likely upcoming referents. Similarly, Mitsugi and MacWhinney (2015) showed that L1 English learners of Japanese did not exploit case-marking information to predict upcoming linguistic information like native Japanese participants did. In contrast, the EEG literature shows that bilinguals can predict syntactic gender in L2 during listening (Foucart et al., 2015) and reading (Foucart et al., 2014), at least when the bilinguals' languages are closely related.

Bilinguals also seem to have difficulty predicting word form information. For example, Ito, Corley and Pickering (2018) used a visual world paradigm to test whether bilinguals pre-activated L2 and L1 word form information when listening in L2. Native English and Japanese-English bilinguals listened to constraining sentences in English (*e.g. The tourists expected rain when the sun went behind the* …), and looked at displays containing either a target object (*cloud*; in Japanese: *Kumo*), a phonological competitor for the target object name in English (*clown*), a phonological competitor for the target object name in Japanese (bear; *kuma*), or an unrelated object (*globe*; *tikyuugi*). Native listeners fixated target objects and English competitors more than distractor objects. Non-native listeners only fixated targets more often than distractors, and not English or Japanese phonological competitors, indicating that they only predicted target word semantics and not word form. This is consistent with EEG reading studies, which have failed to find evidence for prediction of word form in bilinguals in L2 (reading) (Ito, Martin, & Nieuwland, 2017; Martin et al., 2013; but note that recently, Nieuwland et al., 2017, failed to replicate prediction of word-form in native-speakers in a multi-lab study).

Unlike prediction of syntactic and word form information, prediction of semantics is usually not affected by language (L1 vs. L2) (Dijkgraaf et al., 2017; Hopp, 2015; Ito, Corley, et al., 2017). However, Dijkgraaf et al. (submitted) recently showed that differences between prediction of target word semantics in L1 and L2 in bilinguals can also occur when lexical-semantic cues are used for prediction. English-Dutch bilingual participants listened to sentences in L1 and L2 (e.g. '*Her baby doesn't like drinking from a bottle'*) and viewed picture displays while their eye movements were measured. The displays contained an experimental picture (either a target: *bottle,* or a semantic competitor: *glass*) and three unrelated object pictures. Participants were more likely to focus on target objects than on unrelated objects before the auditory target could affect eye gaze. This prediction effect was larger when bilinguals listened to sentences in the L1 than in the L2. Bilinguals were also more likely to look at semantic competitor objects than at unrelated objects, and the bias to competitor objects was larger when the competitor was more strongly related to the target. This relatedness effect was larger in the L1 than in the L2. Dijkgraaf et al. used sentences with variable length, syntactic complexity (e.g. compound sentences), and cloze probability. Also, picture displays were shown only 500 ms before target onset, whereas pictures were shown much earlier in other studies (before sentence onset) (Dijkgraaf et al., 2017; Hopp, 2015; Ito, Corley, et al., 2017), possibly priming the potential referents. These factors may have increased the effort needed to generate predictions in Dijkgraaf et al. and thereby the likelihood of finding language effects. Ito et al. (2018) also found that prediction of semantics was delayed in L2 compared to L1, also indicating that prediction of semantic information can differ between L1 and L2.

Interestingly, there is not that much research on the origin of such differences between prediction in L1 versus L2, when they arise. A first possibility is that they arise because processing L2 taxes working memory more than processing L1 (Abutalebi, 2008; Francis & Gutiérrez, 2012; McDonald, 2006). Perhaps the differences in predictive processing between L1 and L2 found in more demanding paradigms or settings are driven (at least partly) by differences in availability of cognitive resources. If so, an additional reduction in cognitive resources by a load should be especially detrimental for prediction in L2. A second possibility is that prediction differences between L1 and L2 arise from the fact that L2 processing is slower than L1 processing (Cop, Keuleers, Drieghe, & Duyck, 2015; Hahne, 2001; Moreno, Rodríguez-Fornells, & Laine, 2008). Slower processing in L2 may lead to slower use of sentence context to built up higher order meaning and slower subsequent prediction, particularly in fast speech. This hypothesis implies that slowing down speech input in L2 may alleviate the effect of non-native language on prediction.

However, it is important to consider the possibility that load or speed manipulations may not only affect prediction mechanisms directly, but also the balance between different types of prediction mechanisms. In a recent article, Pickering and Gambi (2018) suggest that there are two 'routes' to prediction: one involving the speech production system ('prediction-by-production') and one involving associative mechanisms ('prediction-by-association'). The authors suggest that the prediction-by-production route is the most effective, but also that prediction-by-production is optional, as it requires time and resources. In contrast, prediction-by-association involves spreading activation between concepts stored in long-term memory and it is not optional. According to Pickering and Gambi's proposal, prediction-by-production during comprehension proceeds in the same order as language production: first semantics, then syntax and then word form. Limited availability of resources and slower processing in L2 may cause bilinguals to have difficulty with the prediction-by-production route in L2. The finding that differences between prediction in L1 and L2 have been found most often at the syntactic and word form level, and less so at the semantic level, suggests that this is indeed the case.

Note also that between-language differences are more likely to occur in the auditory, rather than visual modality. Auditory and visual language comprehension differ in that only the latter operates on simultaneously presented information, whereas speech unfolds over time. In addition, speech is highly variable (e.g., due to disfluencies, accent, speech rate, prosody) compared to written language. This may impact effects of processing speed on prediction as listeners cannot control input speed like in (natural) reading, and they need to deal with variability. In addition, spoken language processing may be particularly slow and resource consuming in bilinguals compared to written language processing due to effects of cross-linguistic interference and difficulties distinguishing non-native phoneme contrasts in L2 (Weber & Broersma, 2012). Below, we will discuss evidence for the involvement of cognitive resources and processing speed in prediction during comprehension (mainly from unilingual studies), with a particular focus on the auditory modality. This will be important for the present, bilingual, study.

Cognitive Resources and Prediction

There is some evidence that prediction indeed requires cognitive resources. For example, Huettig and Janse (2016) investigated effects of individual differences on predictive eye-movements in a visual world experiment. Participants listened to sentences such as *Kijk naar de*$_{COM}$ *afgebeelde piano*$_{COM}$ ('look at the displayed piano') while they looked at a four-picture display. Only one of the depicted objects matched the gender of the article in the sentence, so that article gender could be used as a cue for predicting the likely subsequent referent. In addition, participants did multiple tests assessing their working memory capacity, processing speed and non-verbal intelligence. Participants indeed used gender cues to predict the sentence final noun, and participants' working memory capacity and (general) processing speed accounted for most of the variance in anticipatory eye-movements. Huettig and Janse suggested that working memory resources are needed to "ground language in space and time, allowing for short term connections among objects and linking linguistic and visual-spatial representations" (p. 89) (see also Huettig, Olivers, & Hartsuiker, 2011 for a similar proposal). Thus, the more working memory resources are available, the more predictive looking behavior.

In a bilingual study, Ito et al. (2017) tested the assumption that cognitive resources are required for prediction by manipulating cognitive load, using a visual world paradigm. Participants listened to sentences with a predictive or neutral verb while they looked at a display with a target, a competitor from the same semantic category, and two unrelated items. Half of the participants were given a concurrent dual-task (word recall). The authors tested whether the cognitive load had a different effect on prediction in L1 than in L2 speakers of English (with various L1s). Both L1 and L2 speakers looked more at target pictures in the predictive than in the neutral condition, to the same extent. This prediction effect was delayed in the cognitive load condition. Just like the prediction effect, the cognitive load effect did not interact with language, which suggests that L2 speakers use the same mechanisms as L1 speakers to make predictions.

Input Speed and Prediction

The second (possibly related) factor that may modulate predictive language comprehension, and hence differences between L1 and L2, is processing speed. Processing speed may play a role in prediction if it determines the speed of information retrieval from long-term memory, as well as the speed of integration of unfolding information into a representation of sentence meaning (Huettig & Janse, 2016). Consistent with this idea, the monolingual study of Huettig and Janse (Huettig & Janse, 2016) found that besides working memory, general processing speed predicted language mediated anticipatory eye movements as well.

Using a visual world paradigm, Kukona, Fang, Aicher, Chen and Magnuson (2011) provided evidence that prediction (by monolinguals) is indeed enhanced when more processing time is available. Following a predictive verb in simple SVO sentences (e.g. '*Toby arrests the crook'*), participants fixated pictures of verb related agents ('policeman') and patients ('crook') almost to the same extent (the agent role in the sentence was already filled by another entity). Implying that …. When using passive OVS sentences (e.g. *'Toby was arrested by the policeman'*), participant fixated verb-related agents ('policeman') more than related patients ('crook') in the prediction time window (although there were also more fixations to related patients than to distractors). The authors suggest that the difference between the effects found for SVO and OVS sentences could be due to the additional syntactic information in OVS sentences (additional words for passive construction "*by the*"), but also to the additional time available for generating a prediction in this condition. Thus, when enough processing time is available, people engage in active prediction of likely upcoming referents in addition to prediction via passive associative mechanisms. There are also a number of (monolingual) EEG studies on prediction during written language comprehension that support the claim that prediction is enhanced (Chow, Lau, Wang, & Phillips, 2018; Wlotko & Federmeier, 2015) or extended to the word form level (Ito, Corley, Pickering, Martin, & Nieuwland, 2016) when enough time is available.

Present Study

In the present study, we used the visual world paradigm to test whether cognitive load or rather processing speed underlies differences in predictive eye movements between L1 and L2 listening. Most previous research compared prediction between (different) native and non-native participants. As prediction is shaped by experience (Foucart, 2015) and modulated by individual cognitive differences (Huettig & Janse, 2016; Rommers, Meyer, & Huettig, 2015), we believe that it is important to study differences between prediction in L1 and L2 using a within-participants design. Here, participants listened to a semi-naturalistic set of sentences in L1 and L2 with variable cloze probabilities (materials used in Dijkgraaf et al., submitted). Before the onset of target words, participants saw a picture display containing an experimental picture (either a target or a semantic competitor) and three unrelated objects. We investigate the mechanisms underlying the difference between prediction in L1 and L2 in the auditory modality, because we assume that between-languages differences are particularly likely to surface there.

In Experiment 1, cognitive load was manipulated within-participants in a blocked design. Specifically, in half the trials in each language participants were asked to remember 9 (non-word) syllables. This way we tested whether limiting availability of cognitive resources impacted L2 processing more than L1 processing.

Above, we discussed studies that provided evidence that prediction is affected by slower stimulus presentation rate in monolingual language processing. In this view, it is especially interesting whether processing speed may provide an alternative explanation, besides cognitive load, of why prediction may be less pronounced in L2, relative to L1. In Experiment 2, we tested whether speed of processing is a viable explanation for L1/L2 prediction differences in auditory comprehension. We therefore experimentally manipulated stimulus presentation rate (of the same constraining sentences and in the same visual world paradigm as in Experiment 1), and measured the consequences for L2 (and L1) prediction. We asked whether reduced presentation speed makes L2 prediction look like L1, and whether increased speed would do the reverse for L1.

The load and speed manipulations might also shed light on the hypothesis that predictions involves multiple routes. That is to say, a cognitive load and presentation rate may effect target prediction, but not (or to a lesser extent) competitor prediction. A finding like this would be compatible with multi-mechanisms accounts of prediction (Huettig, 2015; Huettig & Janse, 2016; Kuperberg, 2007; Pickering & Gambi, 2018; Pickering & Garrod, 2013), which assume that there are (at least) two routes to prediction: one based on higher level information (using combinatorial, simulation and/or production mechanisms) that likely requires resources, and one based on simple associative connections in which activation spreads automatically between representations. If target prediction is accomplished mostly by the first route and semantic competitor prediction by the second, a cognitive load and presentation rate are expected to have a bigger impact on target prediction than on competitor prediction.

EXPERIMENT 1

Methods

Participants. Seventy-four Dutch-English bilinguals (57 female, age: *M*=20, *SD*=2.2) participated. All participants' dominant and most proficient language was Dutch and English was their second most proficient language. On average, English was acquired from age 11 (*SD*=2.08). The participants had mainly learned English in school, on (social) media, and during holidays abroad. On average, participants reported to encounter Dutch 76% of the time and English 19% of the time. Besides English, participants had knowledge of French (64 participants), German (40 participants), Spanish (10 participants), Portuguese (1 participant), Hungarian (1 participants), Italian (1 participant), and Armenian (1 participant). All participants had normal hearing and normal or corrected-to-normal vision and none of the participants had a language disorder. The bilinguals' proficiency scores are reported in Table 1.

Table 1

Participants' Mean (SD) L1 and L2 LexTALE Scores and Self-ratings

	L1 dutch	L2 english	p-value [c]	df
Lextale[a]	88.78 (8.27)	70.91 (10.91)	<.001	73
Rating listening[b]	4.96 (.20)	4.07 (.56)	<.001	72
Rating speaking[b]	4.96(.20)	3.53 (.62)	<.001	72
Rating reading[b]	4.96(.20)	3.91 (.55)	<.001	72
Rating general[b] proficiency	4.95 (.23)	3.69 (.54)	<.001	71

[a] Scores consist of percentage correct, corrected for unequal proportion of words and nonwords (Lemhöfer & Broersma, 2012).

[b] Means of self-assessed ratings on a scale of 1 to 5 (1=not at all, 5=perfect/mother tongue) for listening, speaking, reading and general proficiency.

[c] Reported p-values indicate significance levels of dependent samples t-tests between scores for Dutch and English in bilinguals. Degrees of freedom differ because 2 participants failed to provide some of the self-ratings.

Materials and Design. The stimulus set for the experiment consisted of 362 Dutch sentences and their English translation equivalents. These sentences were selected out of an initial set of 871 candidate sentences, based on the results of a cloze test filled out by 50 participants (25 in Dutch and 25 in English) that did not participate in the main experiment. The English candidate sentences were checked for grammaticality by a native speaker. For each sentence, the target word with the highest cloze probability was selected. Candidate sentences were included in the final stimulus set if the sentence final word provided by the participants were translation equivalents and if a picture of the word was available in the normed picture set of Severens, Van Lommel, Ratinckx and Hartsuiker (2005). Also, for each provided target word/picture (and translation equivalent), only one sentence was included. Mean cloze probabilities were .71 (*SD*=.23) in Dutch and .68 (*SD*=.24) in English.

All final sentences were paired with five pictures from the Severens et al. (2005) picture set: A target (depicting the sentence final word), a semantic competitor, and three pictures of objects unrelated to the target word. With each auditorily presented sentence the participants saw a four-picture display with either the target picture or the competitor picture (counterbalanced across participants), and three unrelated pictures. Each target picture was also presented in each other position (as competitor and three times as unrelated picture) with different sentences. The English and Dutch translation equivalents of the picture names were phonetically dissimilar (normalized phonetic Levenshtein distance *M*=.25, *SD*=.25). English and Dutch are related languages with many cognates (Schepens, Dijkstra, Grootjen, & van Heuven, 2013). To approximate the distribution of words in natural language, cognates were included in our materials. However, as word activation may be affected by the phonetic similarity between a word and its translation equivalent, we included normalized phonetic Levenshtein distance as a factor in our analyses.

The SNAUT database (Mandera, Keuleers, & Brysbaert, 2017) was used to determine semantic distances between targets and competitors. The semantic distances between each target and competitor were always smaller than .8 on a scale from 0 to 1, with smaller distances for more related word pairs), while the semantic distance between targets and unrelated picture names was always more than .8. The cut-off point of .8 was the lowest point for which it was possible to pair each target word with a competitor word from the same pool. Target and competitor words did not start with the same phoneme (except for one pair in Dutch, *orange-lemon*, *sinaasappel-citroen*).

Forty further Dutch-English bilinguals (from the same participant pool) provided plausibility ratings for target words, competitor words and an unrelated word as sentence endings (20 participants for English sentences and 20 for Dutch sentences)[2] Plausibility ratings did not affect competitor selection, but they were included as factor in the analyses. An example stimulus set and detailed information on our stimuli (cloze probability, stimulus word frequency, word length, competitor plausibility, semantic distance) can be found in Dijkgraaf et al. (2018). Table A1 in Appendix 3A contains the sentences and object names of the target and competitor pictures for each stimulus set.

To ensure that participants continued to pay attention to the sentences, a simple yes/no question about the preceding sentence was presented visually each 12 sentences. Each question was followed by a filler sentence selected from the same initial sentence set as the experimental sentences. The targets of the filler sentences were never included in the visual display, and the filler displays never included the same pictures as experimental sentence displays. The same fillers were presented in each language (mean cloze probability Dutch=.64, mean cloze probability English=.57) (unlike the experimental sentences).

Participants were given an additional cognitive load during half of the English and half of the Dutch trials (load/non-load trials were blocked and order was counterbalanced between participants). Thus, we had a 2 (Language: Dutch vs. English) x 2 (Load: 0 vs. 9 syllables) x 2 (Item type: Target/Competitor vs. Unrelated) design, with all variables manipulated within-subjects.

Recordings. The sentences were recorded by a Dutch-English bilingual (female, 21 yrs) from Flanders. She lived in England from age five to twelve. Her accent was rated by the participants in the experiment a 3.6 in English and 4.4 in Dutch on a scale from 1 (very foreign accent) to 5 (native accent). The sentences were pronounced at a relaxed but natural rate (on average 220 words per minute). The clearest pronunciation of three recordings (sampling frequency 48 kHz) was selected for the experiment. Praat (Broersma & Weenink, 2014) was used to mark target onsets in each sentence and to create versions of each recording in which presentation rates of the sentences were manipulated. The average target word length in the non-manipulated recordings was 507 ms (range 224-942 ms) in English and 511 (240-1168 ms) in Dutch. On average, the sentence leading up to

[2] Due to an error in the test plausibility ratings for three (out of 724 sentences) were missing.

the target word was 1977 ms in English (range 708-4557 ms) and 2164 ms in Dutch (range 764-4764 ms).

Procedure. Participants received written and oral instructions to listen carefully to Dutch and English sentences and to look at pictures on the screen. They were instructed to look wherever they wanted, but not to let their gaze leave the screen (Huettig & Altmann, 2005; McQueen & Huettig, 2012). In addition, participants were asked to answer the occasional yes/no question about a preceding sentence by pressing "j" for yes and "f" for no. Right eye movements were recorded with an Eyelink 1000 eye-tracker (SR Research) (1000 Hz) in tower mount.

Presentation of the auditory sentence over headphones was preceded by a fixation cross on screen for 500 ms. Following Rommers et al. (2013), pictures were presented only 500 ms before the onset of the auditory target word in the sentence. This way, we strived to avoid visual priming of target or competitor word semantics by the visual target or competitor. Picture location was randomized. The picture display remained on the screen for 1000 ms after the auditory sentence had ended. A drift check was performed before proceeding to the next trial.

The sentence pairs (where one sentence's target was the other sentence's competitor and vice versa) were split into two lists (list A and list B). Each sentence could be presented with a target or a competitor picture and each sentence could be presented in Dutch and in English. The participants were presented with one block of a list with 181 sentences (and 12 fillers) in Dutch and one block of the other list with 181 sentences (and 12 fillers) in English. The English and Dutch blocks were subdivided into a load block and a non-load block. In the load blocks, participants were first presented (visually) with 9 non-word syllables of two or three letters. The syllables (from a set of 144 in total) occurred in at least ten words in English and in Dutch and they were matched for frequency between languages (Dutch M=730 per million, English M=402 per million, $t(143)=1.48$, $p=.14$).[3] Participants were asked to study the syllables for 30 seconds and to try to remember as many as possible before the experiment continued. Participants were also instructed to keep on listening to the experimental sentences attentively. After the 30 seconds of syllable study time the experiment continued; experimental sentences and picture displays were presented. After twelve trials, an answer screen appeared and participants were asked to type all the syllables they

[3] Syllable frequencies were based on summated LEMMA frequencies per million of the words containing the syllable, extracted from CELEX (Baayen, Piepenbrock, & Gulikers, 1995). The matched syllables were randomly divided into 16 sets of 9 syllables.

remembered. After the answer screen participants were given thirty seconds to study a new set of nine syllables and so on until the end of the load block.

Language order, list (A or B), load block order, and condition (target or competitor) were counterbalanced across participants, resulting in 16 presentation lists with a fixed random sentence order. Calibration was performed before starting each experimental block. The eye-tracking experiment took approximately 75 minutes.

Afterwards, participants completed a digit span task, measuring recall of digit sequences (part of the Wechsler Adult Intelligence Test, 1997). LexTALE Dutch, LexTALE English (Lemhöfer & Broersma, 2012) (see Table 1 for results), and a language background questionnaire. Completion of the additional tests took approximately 20 minutes.

Analyses. The data were analyzed with linear mixed effects models in R (3.4.0) (R Core Team, 2013). The dependent variable was the empirical logit (a quasi-logit transformation suitable for probabilities near 0 or 1) of the proportion of eye-data samples in which there was a fixation to a picture over the total number of samples (Barr, 2008). The samples in which there was a fixation to one of the three unrelated pictures were averaged. We ran separate analyses for the trials with a target in the display and trials with a competitor in the display.

We first analyzed the data of the full time frame in which predictive looking behavior was expected, without taking into account the time course. Approximately 200 ms (Matin, Shao, & Boff, 1993; Saslow, 1967) is needed for planning and executing a saccade. Therefore, the prediction time frame included the eye-data samples starting from 200 ms after the onset of the picture display, to 200 ms after target onset. All continuous predictors were scaled and centered. First, a full model including all theoretically relevant fixed effects and interactions was fitted for the prediction time frame (Table B2 in appendix 4A). The model included random intercepts of participant and item. The main experimental factors and their interactions were always included in the model. Which secondary factors (less relevant to the main goals of the experiment) were included in the model was determined with a backward fitting procedure (the interaction with the smallest t-value was excluded first). Then, a forward fitting procedure was used to determine the random slopes, followed by backward fitting fixed effects again to find the final model (Barr, Levy, Scheepers, & Tily, 2013). A term was excluded if a Chi-square test comparing the model with and without the term was not significant. We report the results for the final model. The p-values for the fixed effects in our models were obtained with lmerTest (version 2.0-33) (Satterthwaite degrees of freedom approximation) (Kuznetsova, Brockhoff, & Christensen, 2016). Post-hoc contrasts were performed with lsmeans (Lenth, 2016) (Kenward-Roger's approximation to degrees of freedom).The final competitor and target models were then used for a time course analysis, in which we fitted the models for each 50ms time bin in the prediction and target time frame (200 ms after display onset up to 500 ms after target word onset). The data and scripts used for the analyses are available online at Open Science Framework (osf.io/8t76r).

Results

Memory task and digit span. On average, participants remembered 4.5 (*SD*=1.3) syllables during the Dutch (L1) trials and 4.5 (*SD*=1.3) during English trials.[4] Mean digit span forward score was 9.4 (*SD*=1.9), indicating that participants remembered sequences of about 6 digits on average.

No-load trials. First, we tested whether we could replicate the smaller prediction effect in L2 than L1 found in Dijkgraaf et al. (submitted) with the same stimuli. To this end, we determined the optimal model for the data of the subset of trials without a cognitive load. For the target data set, there was a main effect of image type (target vs unrelated). Participants fixated the target image more than unrelated images before the auditory target word could affect fixations (β = -1.46, *SE* = .13, *t* = -11.14, *p* <.001). The effect of image type was smaller in L2 than in L1 (β = .43, *SE* = .09, *t* = 4.88, *p* <.001).

In the competitor trials, there was a main effect of image type (β = -.51, *SE* = .11, *t* = -4.57, *p* <.001) as well. However, image type did not interact with language (β = .14, *SE* = .09, *t* = 1.55, *p* = .12) or with semantic distance (β = .10, *SE* = .08, *t* = 1.15, *p* = .25). The threeway interaction between image type, language and semantic distance was not significant either (β = .04, *SE* = .09, *t* = .39, *p* = .70), unlike our findings in Dijkgraaf et al. (2018).

Target trials. Figure 1a shows the time course of the difference between fixation probability to targets and to unrelated images. Table B1 in appendix 4B contains the final model for the target dataset. There was a main effect of image type (target vs. unrelated) (β = -1.48. *SE* = .12. *t* = -12.14. *p* <.001), indicating that overall there were more fixations on target objects than on unrelated objects in the prediction time frame. The image type effect was smaller in the cognitive load condition than in the no load condition (β = .46, *SE* = .09, *t* = 5.30, *p* <.001). The effect of image type was also smaller in L2 than in L1 (β = .47, *SE* = .09, *t* = 5.43, *p* =<.001). Furthermore, the effect of cognitive load on the target fixation bias was also smaller in L2 (English) than in L1 (Dutch) (β = -.27, *SE* = .12, *t* = -2.22, *p* = .026) (Figure 2).

[4] Due to technical problems the average number of remembered syllables was missing for 3 participants.

Even though digit span forward score did not affect looking behaviour directly (there was no significant interaction between digit span forward score and image type, β = -.03, *SE* = .08, *t* = -.39, *p* = .70), participants with a higher digit span showed a larger effect of cognitive load on the fixation bias for the target image (β = .23, *SE* = .06, *t* = 3.65, *p* = <.001).

Phonetic similarity (standardized phonetic Levenshtein distance) of the translation equivalents affected looking behavior differentially in L1 and L2 trials, as demonstrated by the significant interaction between Language, Image type, and phonetic similarity (β = -.18, *SE* = .06, *t* = 3.65, *p* = .001). There was a stronger fixation bias towards target words with a phonetically more similar translation equivalent in L2 sentences (β = .24, *SE* = .09, *t* = 2.53, *p* = .01), but not in L1 sentences (β = .06, *SE* = .10, *t* = .61, *p* = .54). The bias towards the target was weaker if the target picture had already occurred more often in the experiment on other trials (as unrelated image) (β = .13, *SE* = .03, *t* = 4.10, *p* <.001).

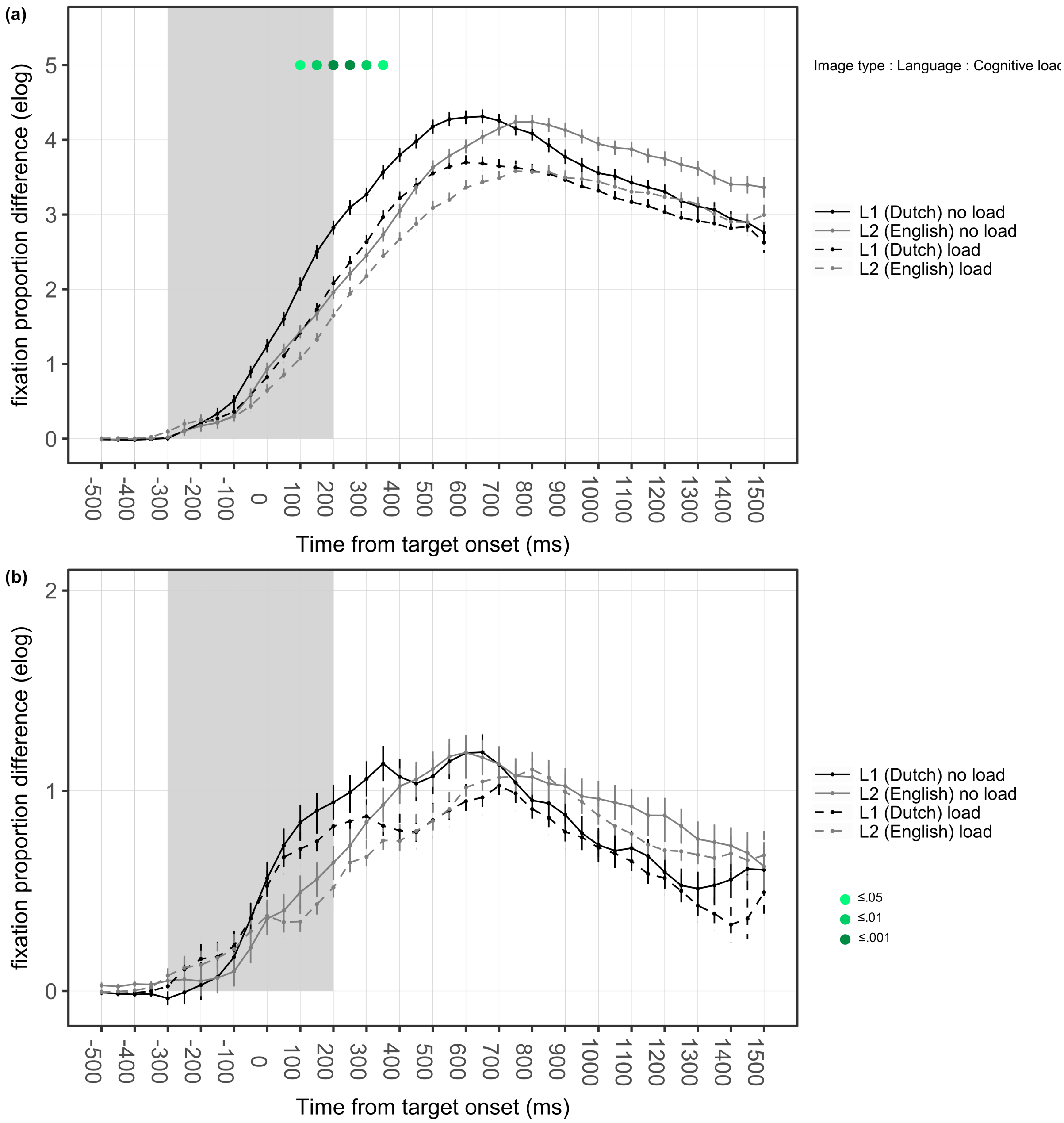


Figure 1a-b. Time course of fixation proportion difference between experimental picture and unrelated pictures. The area shaded in grey indicates the prediction time frame. Whiskers indicate standard errors. Green dots indicate bins where the interaction between image type, language and cognitive load is significant. 1a. Target trials. 1b. Competitor trials

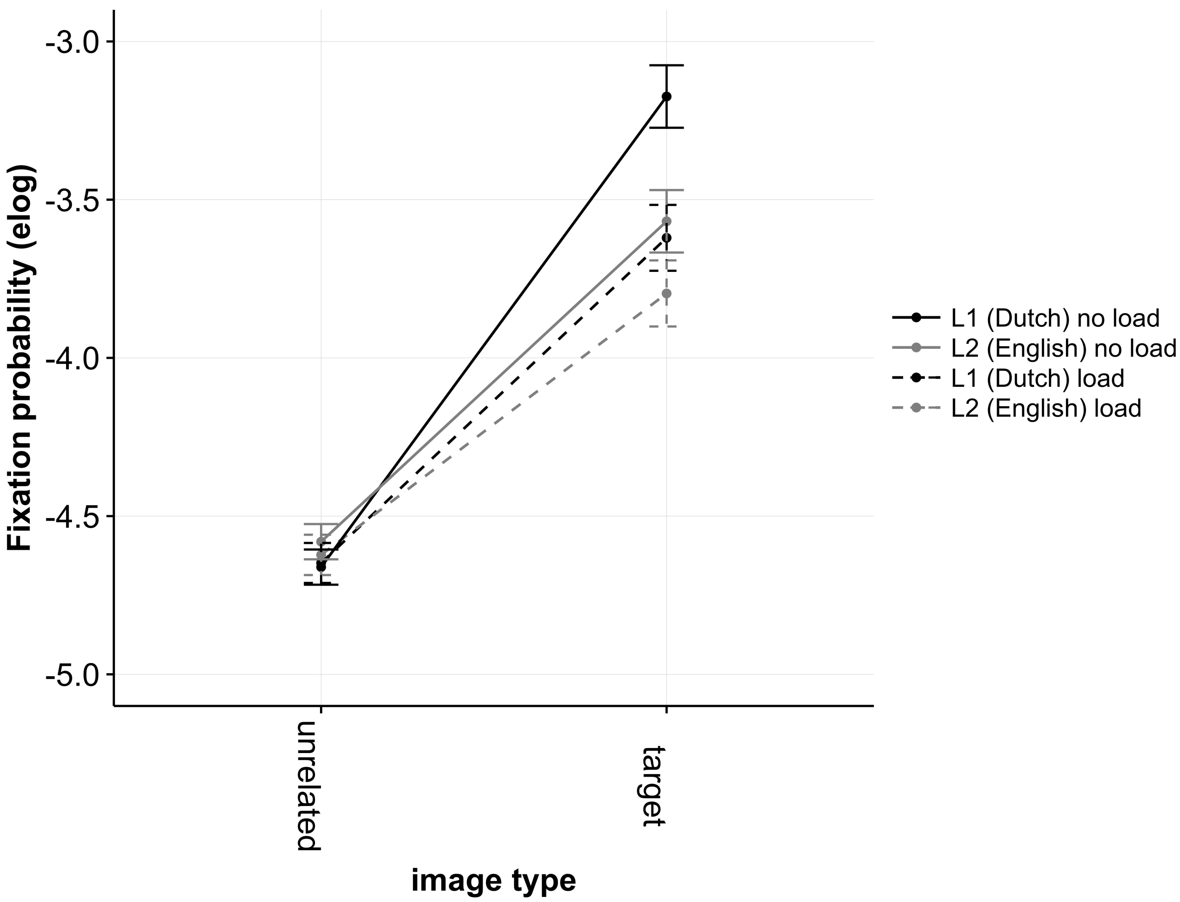


Figure 2. The effect of image type (target vs. unrelated) on fixation probability in each load and language condition. Whiskers indicate standard errors.

Time bin analysis target trials. There were more looks to target objects than to unrelated objects from 200 ms before target onset onwards (Figure 1a). The interactions between image type and load and between image type and language became significant 100 ms later. The three-way interaction between language, image type and load was significant for six consecutive time bins from 100 ms after target onset until 350 ms after target word onset (See Table B2 in Appendix 4B for a full overview of the time course results).

Competitor trials. Figure 1b shows the time course of the difference between fixation probability to competitors and to unrelated images. Table B3 in Appendix 4B contains the final model for the competitor dataset. There was a main effect of image type (competitor vs. unrelated) (β = -.51, *SE* = .11, *t* = -4.59, *p* < .001), indicating that participants looked more at competitors than at unrelated objects. There was also a significant interaction between image type and language (β = .19, *SE* = .09, *t* = 2.22, *p* = .03). The bias towards the competitor object was larger in L1 than in L2. The interaction between image type and load did not reach significance (β = .02, *SE* = .09, *t* = .24, *p* = .81). No other interactions with image type reached significance either (all *p-values*>.1).

Time bin analysis competitor trials. There was a significant bias towards the competitor image from 50 ms before the onset of the target word (Table B4 in Appendix 4B). This effect was modulated by language from time bin 0. The interaction between image type and semantic distance also became significant within the prediction time frame (from 50 ms after target onset). No other factors modulated the bias towards the competitor within the prediction time frame.

Discussion

Experiment 1 tested whether a cognitive load affected predictive eye-movements in L1 and L2. Significant prediction effects were found in both languages, but these effects were stronger in L1 than in L2, like in Dijkgraaf et al. (2018). In the no-load condition only, we also replicated the effect of language on target word prediction during comprehension found in Dijkgraaf et al., using the same paradigm and materials. Prediction of competitors was not modulated by language (unlike Dijkgraaf et al) in the no-load condition only. However, this interaction between language and image type was replicated for competitors across both load conditions (implying the same amount of trials instead of half the amount in the no-load condition), with weaker semantic prediction in L2. Semantic distance between targets and competitors did not modulate prediction until 50ms after target onset, and this effect did not interact with language, like it did in Dijkgraaf et al.

As for our cognitive load manipulation, prediction effects were weaker under a cognitive load in both languages. Our results generalize monolingual and bilingual work showing that resources are required for prediction (Huettig & Janse, 2016; Ito, Corley, et al., 2017). Contrary to our expectations however, the effect of cognitive load was larger in L1 than in L2. One possible explanation for this finding could be related to the relative emphasis on resource intensive prediction-by-production, and ‘passive’ prediction-by-association (Pickering & Gambi, 2018). Specifically, bilinguals may use the prediction-by-production route less in L2 than in L1, because they lack the required time and resources for this prediction mechanism, when listening to L2 speech.

We expected participants with larger working memory capacity to show weaker effects of cognitive load, as they should have more spare resources available for prediction. Surprisingly, participants with a higher digit span forward score showed a larger effect of load on prediction. This is again consistent with multiple mechanism accounts of prediction (e.g., Huettig, 2015; Kuperberg, 2007; Pickering & Gambi, 2018; Pickering & Garrod, 2013) in which one mechanism (for instance, prediction-by-production) is resource intensive. People with larger working memory capacities may use this route more and therefore the effect of an additional load is also larger in those participants. An alternative explanation could be that participants with larger working memory capacity have put more effort in rehearsing the syllables in the load condition, at the expense of listening to the sentences. However, this option was not supported by the correlation between working memory capacity and average recall ($r=.18$, $n=72$, $p=.13$).

If L2 processing indeed taxes working memory more than L1 processing, weaker performance on the secondary (memory) task would be expected during L2 trials. But this was not the case: participants' performance on the working memory task was highly similar during English and Dutch blocks. Thus, it seemed like an equal amount of resources was reserved for the memory task in L1 and L2.

EXPERIMENT 2

Experiment 2 tested whether processing speed may offer an alternative explanation, besides cognitive load, why predictive processing may differ between L1 and L2, by manipulating auditory stimulus presentation rate in both L1 and L2.

Method

Participants. Seventy-five further Dutch-English bilinguals (57 female, age: M=20, SD=3.1) took part in the experiment. All participants' dominant and most proficient language was Dutch and English was their second most proficient language. On average, English was acquired from age 11 (SD=3.2). The participants had mainly learned English in school, on (social) media, and during holidays abroad. On average, participants reported to encounter Dutch 74% of the time and English 21% of the time. Besides English, participants had knowledge of French (72 participants), German (43 participants), Spanish (13 participants), Swedish (1 participant), Afrikaans (2 participants), Italian (1 participant), and Arabic (1 participant). All participants had normal hearing and normal or corrected-to-normal vision and none of the participants had a language disorder. L1 and L2 proficiency was again assessed with the LexTALE vocabulary knowledge test and with self-ratings. The bilinguals' proficiency scores are reported in Table 5.

Table 5

Participants' Mean (SD) L1 and L2 LexTALE Scores and Self-ratings

	L1 dutch	L2 english	p-value [c]	Df
Lextale[a]	89.92 (6.36)	73.67 (10.21)	<.001	74
Rating listening[b]	4.99 (.12)	4.10 (.53)	<.001	73
Rating speaking[b]	4.97(.16)	3.59 (.70)	<.001	74
Rating reading[b]	4.97(.16)	4.05 (.60)	<.001	74
Rating general[b] proficiency	4.96 (.20)	3.72 (.63)	<.001	73

[a] Scores consist of percentage correct, corrected for unequal proportion of words and nonwords (Lemhöfer & Broersma, 2012).

[b] Means of self-assessed ratings on a scale of 1 to 5 (1=not at all, 5=perfect/mother tongue) for listening, speaking, reading and general proficiency.

[c] Reported p-values indicate significance levels of dependent samples t-tests between scores for Dutch and English in bilinguals. Degrees of freedom vary because two participants did not provide all of the solicited ratings.

Materials and Design. The materials and design were the same as in Experiment 1, except that half of the recordings were sped up in Dutch (L1) and half of the sentence recordings were slowed down in English (L2). For the manipulation of the recordings, Dutch sentences were sped up (factor .78) and English sentences were slowed down (factor 1.22) using the "Lengthen (overlap-add)" function in Praat. The recordings were also scaled so that each would have the same peak amplitude (.99).

The used speed factor was based on the results of Dijkgraaf et al. (2018), who used the same stimuli and design, in the same population of bilinguals, but without manipulation of presentation rate. To calculate the required relative delay in lexical activation in L2, we compared the timing of maximum target word activation in each language. The maximum target activation (determined by the first time bin where fixation probability was not significantly higher than the previous) occurred 450 ms after target onset in Dutch and 100 ms later (550 ms after target onset) in English. With a 100 ms delay over 450 ms, activation was 22% slower in English than in Dutch, hence the speed manipulation factor for the present study. Examples of the resulting manipulated recordings can be found online at Open Science Framework (osf.io/8t76r) . None of the participants mentioned presentation speed when asked whether they noticed anything about the sentences. When asked directly whether they noticed that some sentences were fast and some slow, however, almost all participants indicated that they did notice variation in speech rate (particularly the fast version), but that the sentences were nevertheless comprehensible.

Procedure. Half of the sentences presented in each language block were manipulated (speeded up in L1 and slowed down in L2). The manipulated sentences were interspersed between the non-manipulated sentences. There were two fixed random sequences of normal and manipulated sentences. Language order, list (A or B), manipulated/non-manipulated fixed random sequence, and condition (target or competitor) were counterbalanced, resulting in 16 presentation lists with a fixed random sentence order. Calibration was performed before starting each experimental block. The eye-tracking experiment took approximately one hour.

Results

Digit span. Mean digit span forward score was 9.5 (*SD*=1.7).

Non-manipulated trials. First, we tested whether we could replicate the L2 disadvantage in predictive processing found by Dijkgraaf et al. (submitted) and in Experiment 1. To this end we determined the optimal model for the data of the subset of trials that were not manipulated (normal speed). For the target data set, there was a main effect of image type (target vs. unrelated). Participants fixated the target image more than unrelated images before the auditory target word could affect fixations (β = -1.52, *SE* = .13, *t* = -11.89, *p* <.001). The effect of image type was smaller in L2 than in L1 (β = .24, *SE* = .09, *t* = 2.84, *p* = .005).

In the competitor dataset, there was a main effect of image type (β = -.35, *SE* = .12, *t* = -2.86, *p* = .004) as well. Like in Experiment 1, image type did not interact with language (β = .08, *SE* = .09, *t* = .87, *p* = .39) or with semantic distance (β = .09. *SE* = .08. *t* = 1.02. *p* = .31). The threeway interaction between image type, language and semantic distance was not significant either (β =- .06, SE = .09, t = -.72, p = .47).

Target trials. Figure 3a shows the time course of the difference between fixation probability to targets and to unrelated images in each speed condition and language. Note that the manipulated conditions were the fast condition in Dutch and the slow condition in English. Table B5 in Appendix 4B contains the final model for the trials on which the display contained a target picture. The threeway interaction between Language (L1 Dutch vs. L2 English), Image type (target vs. unrelated) and speed (fast vs. slow) did not reach conventional levels of significance (β = -.24, SE = .12, t = 1.95, p = .052). The two-way interaction between Image type and Speed did reach significance (β = .23, SE = .09, t = -2.65, p = .008). Post hoc tests showed that the effect was mainly driven by a difference between fast and slow sentences in L1 (β = .23, SE = .09, t =-2.65, p = .008). The difference between normal and slowed down sentences in English did not reach significance (β = .008, SE = .09, t = .10, p = .92) (Figure 4).

There was a stronger bias towards targets that had phonetically more similar translation equivalents, and this effect was larger in Dutch than in English (β = .16, SE = .06, t = 2.65, p = .008). The two-way interaction between image type and phonetic similarity however, did not reach significance (β = -.15, SE = .09, t = -1.66, p = .097). Further, the bias towards the target was stronger for sentences with more syllables (β = -.23, SE = .07, t = -3.23, p = .001), sentences with a higher cloze probability (β = -.17. SE = .06. t = -2.93. p = .003), and sentences in the second block compared to the first block (β = -.22, SE = .09, t = -2.64, p = .008). The bias towards the target was weaker if the target picture had already occurred more often in the experiment in other positions (as unrelated image) (β = .09, SE = .04, t = 2.10, p = .04). English proficiency (lexTALE score) and the interactions between English proficiency, language and image type did not affect looking behavior ($\chi(4)$=2.20, p=.70) and they were therefore not included in the final model.

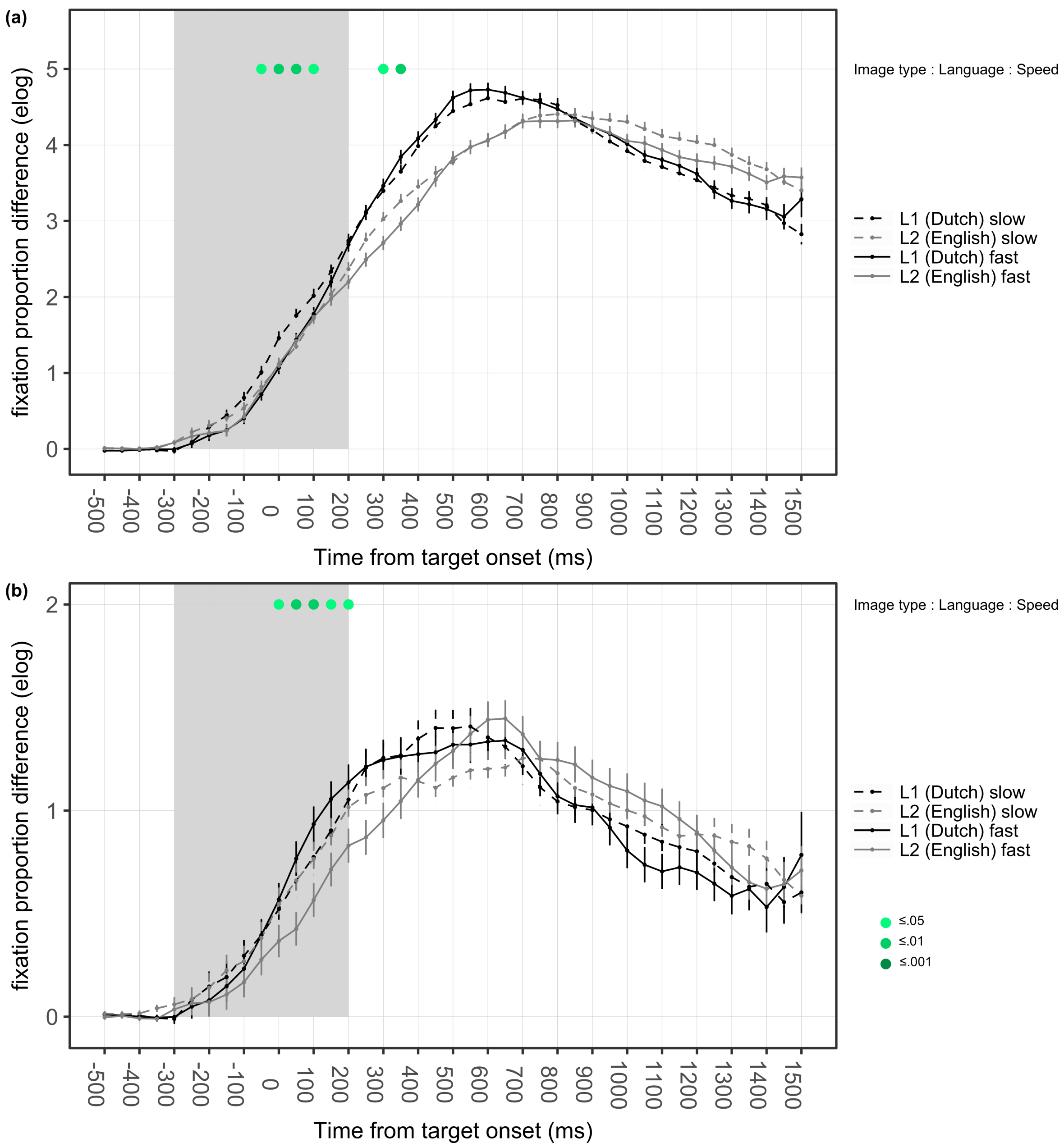


Figure 3a-b. Time course of fixation proportion difference between experimental picture and unrelated pictures. The area shaded in grey indicates the prediction time frame. Whiskers indicate standard errors. Green dots indicate bins where the interaction between image type, language and cognitive load is significant. 1a. Target trials. 1b. Competitor trials

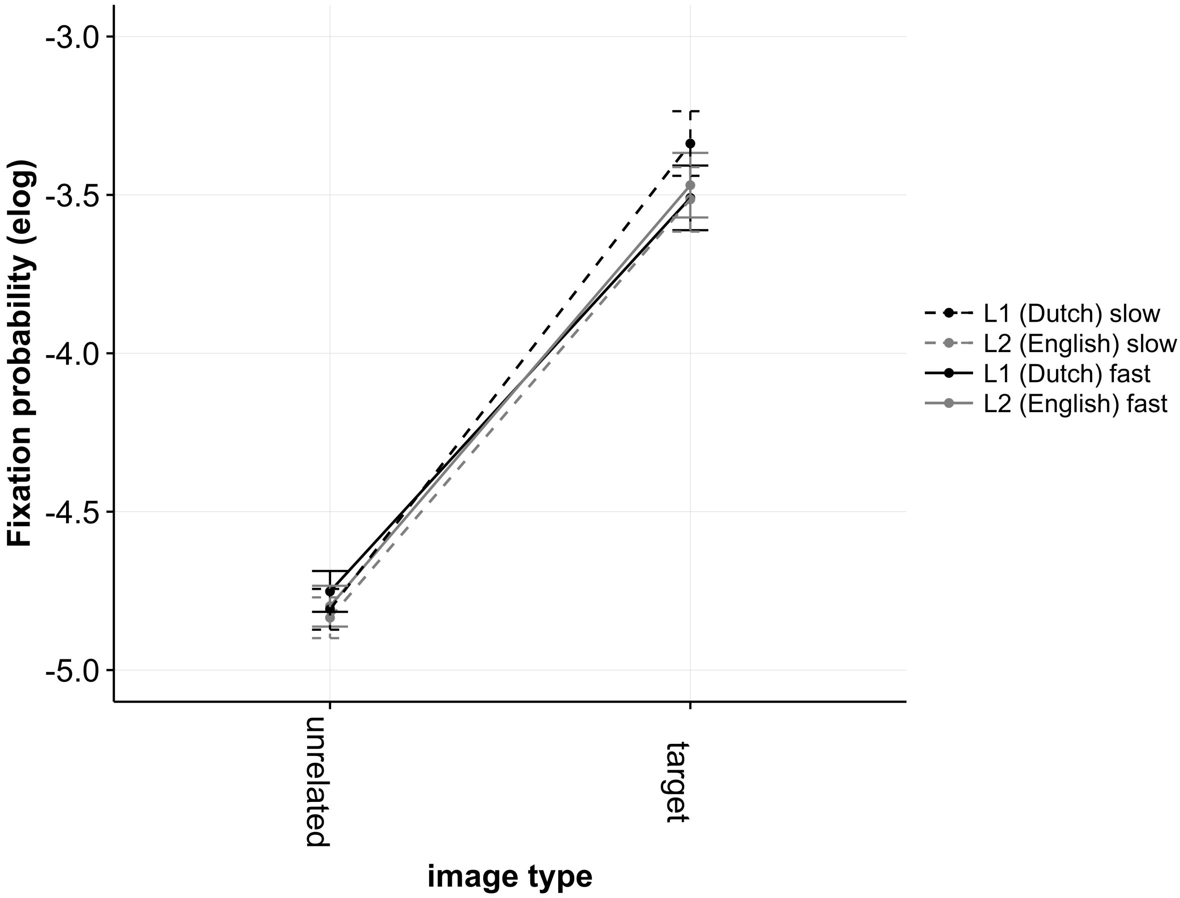


Figure 4. Effect of image type (unrelated vs. target) on fixation probability in each language and speed condition. Whiskers indicate standard error.

Time bin analysis target trials. The main effect of image type became significant 150 ms before target onset (see Table B6 in Appendix 4B). The interaction between image type and language did not reach significance until after the prediction time frame (200-500 ms). The interaction between Language, Image type, and Speed manipulation reached significance in the bins -50, 0, 50, and 100 ms (see Figure 3a). In these bins the bias towards the target was largest in the L1-slow condition, followed by L2 fast and L2 slow. The bias towards the target was weakest in the L1 fast condition. Post-hoc tests reveal that bias towards the target differed between the L1-slow and L1-fast condition in each of the four time bins (*p-values*<.05). At 0, 50 and 100 ms, the target bias was also larger in L1-slow than in L2-fast and L2-slow (*p-values*<.05).

The interaction between language and image type was only significant from 200 to 500 ms. In these time bins there was a larger image type effect in L1 (Dutch) than in L2 (English), and speed condition no longer affected fixation bias to targets.

Competitor trials. Figure 3b shows the time course of the difference between fixation probability to competitors and to unrelated images in each speed condition and language. Table B7 in Appendix 4B contains the final model for the competitor dataset. There was a significant bias towards competitor images compared to unrelated images (β = -.54, *SE* = .11, *t* = -4.73. *p* <.001). The two-way interaction between Image type and speed condition did not reach significance (β = .09, *SE* = .08, *t* = 1.09, *p* = .27), but the two-way interaction between language and image type (β = .20, *SE* = .09, *t* = 2.20, *p* = .028) was significant, with the fixation bias being stronger in L1 than in L2. Interestingly, the threeway interaction between Language, Image type, and Speed manipulation also reached significance (β = -.27. *SE* = .12. *t* = -2.28. *p* = .02) (Figure 5). The fixation bias towards competitors was weaker in the fast than in the slow condition, but only in L2 (L2: β = -.18, *SE* = .08, *t* = -2.13, *p* = .03, L1: β = .09, *SE* = .08, *t* = 1.10, *p* = .27). The interaction between semantic distance (between the target and competitor) and image type (competitor vs. unrelated) did not reach significance (β = .10, *SE* = .08, *t* = 1.33, *p* = .18), nor did the fourway interaction between Language, image type, semantic distance, and speed reach significance (β =.04, *SE* = .12, *t* = .32, *p* = .75).

The bias towards the competitor was stronger for sentences with more syllables (β = -.22, *SE* = .07, *t* = -3.09, *p* = .002), sentences in the second block (β = -.18, *SE* = .0,. *t* = -2.20, *p* = .03), and competitors with a higher plausibility rating (β = -.21, *SE* = .07, *t* = -3.01, *p* = .003). As in the target data set, the bias towards the competitor was weaker if the competitor picture had already occurred in the experiment in other positions (as unrelated image) (β = .11, *SE* = .04, *t* = 2.44, *p* = .01).

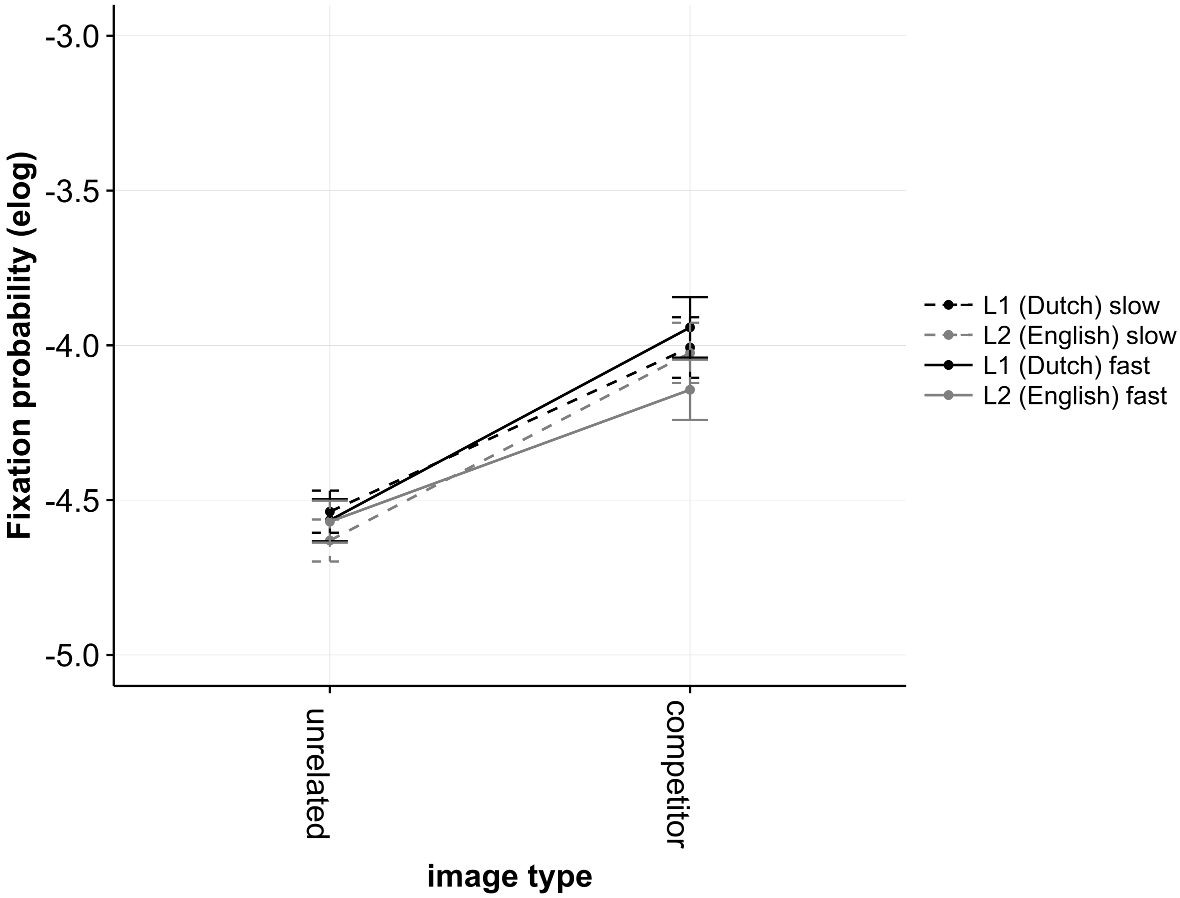


Figure 5. Effect of image type (unrelated vs. competitor) on fixation probability in each language and speed condition. Whiskers indicate standard error.

Time bin analysis competitor trials. The main effect of image type became significant 50 ms before target word onset and remained significant in all further time bins (Table B in Appendix 4B). The threeway interaction between language, image type, and speed condition was significant from 0 to 200 ms after target onset (Figure 3b). In the 0, 50 and 100ms time bins, there was a weaker competitor bias in the L2-fast condition than in the L2 slow condition (*p-values*<.05). From 50-200ms there was also a weaker bias in the L2 fast condition than in the L1 fast condition. The fourway interaction between image type, language, semantic distance and speed condition did not reach significance in any bin.

The competitor bias was modulated by semantic distance from 100 ms after target onset onwards (the before last time bin falling in the prediction time frame). With a larger image type effect for more related competitors. The effect of semantic distance on the fixation bias to competitors was not affected by language in any of the time bins.

Discussion

Experiment 2 investigated effects of processing speed on predictive processing in L1 and L2. For target trials, the interaction between image type, language and speed almost reached significance. In L1, predictive processing was attenuated when auditory stimulus presentation rate was increased. However, no effect of slower auditory stimulus presentation was found in L2. The time course analysis revealed that the interaction between image type, language and speed manipulation was significant in four consecutive time bins in the prediction time frame. Post-hoc tests showed that this interaction was driven mainly by the difference between the normal vs. speeded L1 conditions, with a weaker prediction effect in the speeded condition than in the normal condition.

For competitor trials, on the contrary, slower stimulus presentation rate enhanced prediction in L2, but faster presentation rate did not attenuate prediction in L1. Even though competitor prediction was enhanced in L2 sentences presented at slower rate, there was still an interaction between image type and language on competitor trials, suggesting that prediction of competitors was weaker in L2 than in L1. Like the results of Experiment 1, the results are compatible with multi-mechanism accounts of prediction, with a different pattern of results for target and competitor pre-activation. Processing speed seems to mainly effect target prediction in L1, and competitor prediction in L2, with slower presentation rate causing competitor pre-activation in L2 to pattern with competitor pre-activation in L1.

The separate analysis on normal speed trials showed a larger target and competitor prediction effect in L1 than in L2, replicating Experiment 1 and Dijkgraaf et al. (2018). Like findings in Dijkgraaf et al., (pre-)activation of normal speed competitors was modulated by semantic distance between competitors and targets but only from the prefinal bin of the prediction time frame). The effect of target-competitor semantic distance on prediction did not differ between languages either.

GENERAL DISCUSSION

The present study compared prediction between L1 and L2 listening. We investigated both cognitive load and processing speed as explanations of weaker L2 prediction, using experimental manipulations of cognitive load and stimulus presentation rate in a visual world paradigm. First, we replicated the finding that target prediction effects were larger in bilinguals' L1 than in L2 twice, consistent with our own previous research with the same stimuli (Dijkgraaf et al., 2018). We also found significant predictive looking behavior to semantic competitors, and the effect of language on competitor prediction was replicated in both experiments across speed and load conditions. These results confirm studies that have shown that under certain circumstances predictive processing is weaker in L2 than in L1 (Dijkgraaf et al., 2018; Hopp, 2015; Ito et al., 2018; Martin et al., 2013; Mitsugi & Macwhinney, 2015; Sagarra & Casillas, 2018; van Bergen & Flecken, 2017). However, such a language difference does not always emerge (Dijkgraaf et al., 2017; Foucart et al., 2015; Hopp, 2015; Ito, Corley, et al., 2017), and the underlying reason for differences remained unknown.

Crucially, we showed that both cognitive load and processing speed had an effect on prediction during listening, and that these factors also interacted with language. A cognitive load attenuated predictive eye-movements to targets in both L1 and L2, but to a stronger extent in L1, and faster stimulus presentation rate did so only in L1. Slower stimulus presentation rate in L2 did not significantly modulate predictive eye-movements to targets. Also, predictive eye-movements to competitors were not affected by cognitive load in either language, but slower presentation speed enhanced prediction of competitors in L2. Below, we discuss the evidence for and against the roles of cognitive load and processing speed on predictive processing in L1 and L2 found in the present study, in relation to previous literature.

Cognitive Load

An additional cognitive load resulted in a weaker bias towards target images in L1 and L2, but contrary to what we expected, the effect of load on the fixation bias toward targets was larger in L1. A cognitive load did not eliminate predictive looking in either language. The finding that prediction effects were weaker when participants were under a cognitive load compliments the unilingual study of Huettig and Janse (2016), which showed that working memory capacity modulates predictive processing. We interpret the finding that a cognitive load interferes with predictive processing as an indication that prediction requires cognitive resources (e.g., Pickering & Gambi, 2018). The findings are consistent with the idea that the brain requires cognitive resources to ground language in space and time (Huettig & Janse, 2016; Huettig et al., 2011); i.e. to link linguistic and visuo-spatial representations stored in long-term memory to the present (or future) context. For example, hearing a word would activate a phonological representation stored in long term memory, as well as its associated semantic and visual representations. Similarly, seeing a picture of an object also activates its visual representation, and its associated semantic and phonological representations, stored in long-term memory. Working memory would enable us to link the activated visual and linguistic information and the resulting activation of the object's visuo-spatial representation will increase the likelihood of (anticipatory) saccadic eye-movements towards that object.

The finding that working memory is involved in predictive processing both in L1 and L2 is also consistent with the idea that predictive processing in L1 and L2 require the same underlying mechanisms (Kaan, 2014). Predictive processing seems to be less efficient in the L2 than in the L1, but an additional cognitive load does not eliminate prediction in L2 altogether. Ito, Corley and Pickering (2017) found that a cognitive load delayed predictive eye-movements in L1 and L2 listeners. However, Ito, Corley and Pickering found no language modulation of the cognitive load effect on prediction, whereas we found that a cognitive load affected L1 prediction more than L2 prediction. One difference between Ito et al.'s (2017) study and the present experiment, that may explain the different finding was the nature of the load task. In Ito et al. participants were asked to remember English words. Thus, the cognitive load may have been inherently heavier on L2 listeners than on L1 listeners: the to be recalled words may have interfered more with the L2 sentence comprehension than with the L1 sentences. In the present experiment, participants were asked to remember non-word syllables that were equally frequent in English and Dutch. This way we strived to make the load conditions more comparable in L1 and L2.

The finding that load had a larger effect in L1 is contrary to our hypothesis that cognitive load would have a greater effect on prediction in L2 because L2 processing is presumably more resource consuming than L1 processing. One explanation could be that non-native listeners rely less on resource intensive prediction-by-production than native speakers (Pickering & Gambi, 2018). This explanation can account for our finding that prediction of targets is weaker in L2 than in L1, but also for the finding that cognitive resources impact L2 prediction less than L1 prediction. Previous literature also provides converging evidence. According to prediction-by-production accounts, prediction proceeds via the same stages as actual production. Thus, prediction of word form is a later stage than prediction of semantics and syntax and should therefore be the first type of prediction to be affected in individuals relying less on prediction-by-production. Indeed, L2 comprehenders do not seem to predict word form (Ito, Martin, et al., 2017; Ito et al., 2018; Martin et al., 2013), like native speakers sometimes do (DeLong, Urbach, & Kutas, 2005; Ito et al., 2016, 2018; Martin et al., 2013). Prediction involving (morpho-)syntactic information is often weaker or slower in L2 as well, (Hopp, 2013, 2015; Kaan et al., 2014). Vice versa, semantic prediction, which is likely to occur via low-level lexical associations (prediction-by-association), is usually intact in L2 comprehenders (Dijkgraaf et al., 2017; Hopp, 2015; Ito, Corley, et al., 2017).

Presentation Rate

Our second manipulation concerned processing speed as an origin of L1 vs. L2 differences. In the present study, speeded presentation of sentences in L1 resulted in weaker predictions of targets. Slower presentation in L2 did not result in stronger predictions of targets, but competitor prediction was enhanced. Our findings partially support the notion that the speed with which information can be retrieved from long-term memory and integrated with contextual information in working memory also affect predictive processing (Huettig & Janse, 2016). Specifically, when native speakers listened to L1 sentences that were speeded up (and therefore had less time available for processing), effects of predictive processing were smaller than when participants listened to sentences presented at a normal rate. On the other hand, if the L2 disadvantage in predictive processing were caused by a processing delay, slower L2 input should result in enhanced target prediction as compared to L2 input at a normal rate. Here, we found no effect of slowing down auditory stimulus presentation rate in L2 on target prediction. This suggests that the prediction disadvantage in L2 is not primarily caused by a processing delay.

If we assume that processing speed indeed indexes the speed with which information can be retrieved from long term memory and the speed with which this information can be linked to other types of information, then there is an obvious link with working memory (assuming retrieval and linking is performed in working memory). Thus, one explanation for the finding that speeded input affects prediction in L1 is that more cognitive resources are required for generating timely predictions based on speeded input. If so, then why was there no benefit of slowing down stimulus presentation rate in L2? One explanation is analogous to the hypothesis we discussed in relation to the larger effect of cognitive load in L1: Predictive processing in L2 may be affected less by processing speed and cognitive load because it relies more on automatic processes ("prediction-by-association", in the framework of Pickering & Gambi, 2018), whereas native speakers rely more on the most effective and most correct "prediction-by-production" route, which requires time and resources. Future research could be aimed at testing this hypothesis directly.

Competitor Prediction

There was no evidence that cognitive load affected prediction of semantic competitors. The null-effect for the cognitive load manipulation is surprising, as load had a strong effect on target prediction. However, the activation of competitors may have come about by different mechanisms than active target pre-activation. Competitor plausibility as sentence ending was relatively low,[5]

[5] (*M*=2.14 *SD*=1.46 on a 7 point scale ranging from 'not likely at all as sentence ending' to 'very likely as sentence ending')

and therefore competitors may have been activated mostly via automatic spreading activation due to low-level semantic associations between targets and competitors, or between words in the sentence and competitors (prediction-by-association). Targets, on the other hand, may have been primarily activated by higher order meaning (based on the combination of words in the sentence) (Huettig, 2015; Kuperberg, 2007). The finding that load affected target prediction and not competitor prediction suggests that cognitive resources are not involved (or involved to a lesser extent) in competitor pre-activation. This supports the idea that multiple mechanisms are involved in prediction during language comprehension (Huettig, 2015; Kuperberg, 2007; Pickering & Gambi, 2018).

Slower stimulus presentation rate enhanced prediction of competitors in L2, but faster rate did not affect competitor prediction in L1. Again, this suggests that listeners may emphasize the prediction-by-production route less in L2 than in L1. If competitor pre-activation is indeed mainly dependent on associative mechanisms, the effect of presentation rate on competitor prediction in L2 suggests that prediction-by-association requires more time in L2 than in L1, perhaps because of weaker associative connections in the less practiced L2 than in L1 (Gollan et al., 2011). This idea however, contrasts with the hypothesis that prediction-by-production, and not prediction-by-association is time and resource consuming.

At first glance, the difference between L1 and L2 competitor pre-activation found in both experiments is unexpected, assuming the hypotheses that target and competitor prediction indeed depend on different mechanisms (prediction-by-production and prediction-by-association, respectively) and that L2 listeners rely less prediction-by-production. However, prediction of competitors in the current experiment could have occurred via associations with words in the sentence as well as via associations with the predicted target words. As target word prediction was likely to depend on prediction-by-production (at least in part), not only prediction of target words but also subsequent automatic spreading activation to competitors should be weaker in L2 than in L1. Peters, Grüter & Borovsky (2015) provide converging evidence, showing that low-proficient non-native listeners rely more on prediction-by-association than high-proficient bilinguals. In their visual world study, low-proficient and high-proficient non-native comprehenders listened to sentences (e.g., *"The pirate chases the ship"*) while they looked at displays featuring agent-related, action-related and unrelated pictures. The low-proficient bilinguals were more likely than high-proficient bilinguals to anticipate locally-coherent action-related distractors (e.g., a cat).

CONCLUSION

Here we investigated whether a difference in availability of cognitive resources and slower processing speed underlie the difference in predictive processing in L1 and L2 listening. This study demonstrated that a cognitive load impacts prediction of targets in the L1 and the L2. Faster speech input impacted prediction of targets in the L1, but slower input did not enhance target prediction in L2. Cognitive load did not affect pre-activation of semantic competitors, suggesting that pre-activation based on low-level lexical associations require less cognitive resources than active prediction of (semantics of) sentence-final target words. Pre-activation of competitors was enhanced by slower input speed in L2, but not in L1. The results are consistent with, and extend multi-mechanism accounts of prediction, developed for monolingual language processing (e.g., Huettig, 2015), and the hypothesis that bilinguals rely less on resource intensive prediction mechanisms in L2 than in L1 comprehension (Pickering & Gambi, 2018).

REFERENCES

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica*, *128*(3), 466–478. https://doi.org/10.1016/j.actpsy.2008.03.014

Baayen, H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database. Release 2 (CD-ROM)*. Linguistic Data Consortium, University of Pennsylvania.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*(4), 457–474. https://doi.org/10.1016/j.jml.2007.09.002

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Broersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer (Version 5.3.78). Retrieved from www.praat.org

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 1029–1040. https://doi.org/10.1037/a0015901

Chow, W.-Y., Lau, E., Wang, S., & Phillips, C. (2018). Wait a second! delayed impact of argument roles on on-line verb prediction. *Language, Cognition and Neuroscience*, *0*(0), 1–26. https://doi.org/10.1080/23273798.2018.1427878

Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015). Frequency effects in monolingual and bilingual natural reading. *Psychonomic Bulletin & Review*, *22*(5), 1216–1234. https://doi.org/oi: 10.3758/s13423-015-0819-2

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121. https://doi.org/10.1038/nn1504

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition*, *20*(5), 917–930. https://doi.org/10.1017/S1366728916000547

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2018). Prediction and integration of semantics during L2 and L1 listening. *Manuscript Submitted for Publication.*

Foucart, A. (2015). Prediction is a question of experience. *Linguistic Approaches to Bilingualism*, *5*(4), 465–469. https://doi.org/10.1075/lab.5.4.04fou

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1461–1469. https://doi.org/10.1037/a0036756

Foucart, A., Ruiz-Tada, E., & Costa, A. (2015). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition*, 1–7. https://doi.org/10.1017/S1366728915000486

Francis, W. S., & Gutiérrez, M. (2012). Bilingual recognition memory: Stronger performance but weaker levels-of-processing effects in the less fluent language. *Memory & Cognition*, *40*(3), 496–503. https://doi.org/10.3758/s13421-011-0163-3

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209. https://doi.org/10.1037/a0022256

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, *30*(3), 251–266. https://doi.org/10.1023/A:1010490917575

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*(1), 33–56. https://doi.org/10.1177/0267658312461803

Hopp, H. (2015). Semantics and Morphosyntax in Predictive L2 Sentence Processing. *International Review of Applied Linguistics*. https://doi.org/10.1515/iral-2015-0014

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, *96*(1), 23–32. https://doi.org/10.1016/j.cognition.2004.10.003

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459

Huettig, F., Olivers, C. N. L., & Hartsuiker, R. J. (2011). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Acta Psychologica*, *137*(2), 138–150. https://doi.org/10.1016/j.actpsy.2010.07.013

Ito, A., Corley, M., & Pickering, M. J. (2017). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, *21*(2), 251–264. https://doi.org/10.1017/S1366728917000050

Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language*, *86*, 157–171. https://doi.org/10.1016/j.jml.2015.10.007

Ito, A., Martin, A. E., & Nieuwland, M. S. (2017). On predicting form and meaning in a second language. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*, *43*(4), 635–652. https://doi.org/10.1037/xlm0000315

Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, *98*, 1–11. https://doi.org/10.1016/j.jml.2017.09.002

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? In *Linguistic Approaches to Bilingualism* (pp. 257–282).

Kaan, E., Kirkham, J., & Wijnen, F. (2014). Prediction and integration in native and second-language processing of elliptical structures. *Bilingualism: Language and Cognition*, 1–18. https://doi.org/10.1017/S1366728914000844

Kukona, A., Fang, S.-Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The Time Course of Anticipatory Constraint Integration. *Cognition*, *119*(1), 23–42. https://doi.org/10.1016/j.cognition.2010.12.002

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Research*, *1146*, 23–49. https://doi.org/10.1016/j.brainres.2006.12.063

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: Tests in Linear Mixed Effects Models*. Retrieved from https://CRAN.R-project.org/package=lmerTest

Lemhöfer, K., & Broersma, M. E. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*(1). https://doi.org/10.18637/jss.v069.i01

Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, *92*, 57–78. https://doi.org/10.1016/j.jml.2016.04.001

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, *53*(4), 372–380. https://doi.org/10.3758/BF03206780

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*(3), 381–401. https://doi.org/10.1016/j.jml.2006.06.006

McQueen, J. M., & Huettig, F. (2012). Changing only the probability that spoken words will be distorted changes how they are recognizeda). *The Journal of the Acoustical Society of America*, *131*(1), 509–517. https://doi.org/10.1121/1.3664087

Mitsugi, S., & Macwhinney, B. (2015). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, *19*(01), 19–35. https://doi.org/10.1017/S1366728914000881

Moreno, E. M., Rodríguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics*, *21*(6), 477–508. https://doi.org/10.1016/j.jneuroling.2008.01.003

Nieuwland, M., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., … Huettig, F. (2017). Limits on prediction in language comprehension: A multi-lab failure to replicate evidence for probabilistic pre-activation of phonology. *BioRxiv*, 111807. https://doi.org/10.1101/111807

Peters, R., Grüter, T., & Borovsky, A. (2015). Anticipatory and Locally Coherent Lexical Activation Varies as a Function of Language Proficiency. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 1865–1870). Austin, TX.

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: a theory and review. *Psychological Bulletin*. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(04), 329–347. https://doi.org/10.1017/S0140525X12001495

R Core Team. (2013). *R: A language and environment for statistical computing.* Vienna, Austria.: R Foundation for Statistical Computing,. Retrieved from http://www.R-project.org/.

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, & Psychophysics*, *77*(3), 720–730. https://doi.org/10.3758/s13414-015-0873-x

Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia*, *51*(3), 437–447. https://doi.org/10.1016/j.neuropsychologia.2012.12.002

Sagarra, N., & Casillas, J. V. (2018). Suprasegmental information cues morphological anticipation during L1/L2 lexical access. *Journal of Second Language Studies*, *1*(1), 31–59. https://doi.org/10.1075/jsls.17026.sag

Saslow, M. G. (1967). Latency for Saccadic Eye Movement*. *JOSA*, *57*(8), 1030–1033. https://doi.org/10.1364/JOSA.57.001030

Schepens, J., Dijkstra, T., Grootjen, F., & van Heuven, W. J. B. (2013). Cross-Language Distributions of High Frequency and Phonetically Similar Cognates. *PLoS ONE*, *8*(5), e63006. https://doi.org/10.1371/journal.pone.0063006

Severens, E., Lommel, S. V., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, *119*(2), 159–187. https://doi.org/10.1016/j.actpsy.2005.01.002

van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language*, *92*, 26–42. https://doi.org/10.1016/j.jml.2016.05.003

Van Berkum, J. J. A. (2010). The brain is a prediction machine that cares about good and bad—any implications for neuropragmatics. *Italian Journal of Linguistics*, *22*(1), 181–208.

WAIS-III: Administration and scoring manual: Wechsler adult intelligence scale. (1997) (Psychological corporation.). San Antonio, TX.

Weber, A., & Broersma, M. E. (2012). Spoken word recognition in second language acquisition. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Hoboken, NJ, USA: John Wiley & Sons, Inc. Retrieved from http://doi.wiley.com/10.1002/9781405198431

Wlotko, E. W., & Federmeier, K. D. (2015). Time for prediction? The effect of presentation rate on predictive sentence comprehension during word-by-word reading. *Cortex*, *68*, 20–32. https://doi.org/10.1016/j.cortex.2015.03.014

CHAPTER 5
IS THERE ADAPTATION OF SPEECH PRODUCTION AFTER SPEECH PERCEPTION IN BILINGUAL INTERACTION?[1]

In dialogue, speakers tend to adapt their speech to the speech of their interlocutor. Adapting speech production to preceding speech input may be particularly relevant for second language (L2) speakers interacting with native (L1) speakers, as adaptation may facilitate L2 learning. Here we asked whether Dutch-English bilinguals adapt pronunciation of the English phonemes /æ/ and coda /b/ when reading aloud sentences after exposure to native English speech. Additionally, we tested whether social context (presence or absence of a native English confederate) and time lag between perception and production of the phoneme affected adaptation. Participants produced more English-like target words that ended in word-final /b/ after exposure to target phonemes produced by a native speaker, but the participants did not change their production of the phoneme /æ/ after exposure to native /æ/. The native English speaking confederate did not show consistent changes in speech production after exposure to target phonemes produced by L2 speakers. These findings are in line with Gambi and Pickering's simulation theory of phonetic imitation (2013).

[1] Broos, W.P.J., Dijkgraaf, A., Van Assche, E., Vander Beken, H., Dirix, N., Lagrou, E., Hartsuiker, R.J., Duyck, W. (in press). Is there adaptation of speech production after speech perception in bilingual interaction? *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

INTRODUCTION

Speech production is highly variable. This variability is caused by between-speaker differences such as the mother tongue, age, gender, dialect, and articulatory properties of a particular speaker. In addition, within-speaker differences manifest themselves through peripheral factors such as the time of day, mood, or even just having a cold. Therefore, articulation of words or even phonemes varies considerably. As a consequence, listeners must find a way to cope with this variation. The fact that listeners mostly do not experience difficulty understanding (variable) speech suggests that they can do this very efficiently. Indeed, studies on speech perception have shown that listeners can quickly adjust their perceptual system, for instance to deal with an unusual way in which a speaker realizes a particular phoneme (e.g., Norris, McQueen, & Cutler, 2003). Such adjustment may be particularly useful in a second language (L2), given that the realization of phonemes varies across languages and that such adjustments may help L2 learning (Costa, Pickering, & Sorace, 2008), especially when interacting with native speakers who master the language better. The goal of the present study is to test whether non-native listeners (of English) are not only sensitive to differences between their own L2 phoneme production and native production, but also whether these differences affect their L2 speech production (in other words, whether there is *alignment* between L2 speech production and perception). To gauge whether any such adaptation is automatic or strategic, we considered the effects of several further variables. First, we tested whether the physical presence of a native speaker has an additional effect on speech alignment, since previous studies suggested that social context modulates alignment (e.g., Babel, 2012). Second, we manipulated the lag (number of intervening trials) between perception and production of the critical phoneme.

Phonetic Alignment in L1 Speech Production

Previous studies have shown that L1 listeners can adjust their perception to speech that is produced by their interlocutor, including accents and other non-native speech characteristics (Bradlow & Bent, 2008; Eisner & McQueen, 2006; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994; Norris et al., 2003). Norris et al. (2003) for instance, demonstrated this by using a paradigm in which participants were exposed to an ambiguous fricative [?], midway between [f] and [s]. When listeners were exposed to ambiguous [f]-final words, they categorized later ambiguous [?] more often as an [f], whereas when listeners were exposed to ambiguous [s]-final words, they categorized the ambiguous [?] more often as an [s]. So, listeners can perform perceptual adaptation by using their lexical knowledge to adjust their phonemic representations, making them consistent with specific speech variants. This effect also occurs when listening in L2 (Weber, Betta, & McQueen, 2014).

There is also evidence suggesting that speakers adapt speech production to speech of an interlocutor. Alignment of speech production occurs at the syntactic (e.g., Bernolet, Hartsuiker, & Pickering, 2012, 2013; Pickering & Branigan, 1999), lexical (e.g., Branigan, Pickering, Pearson, McLean, & Brown, 2011), and phonetic (e.g., Babel, 2012; Lametti, Krol, Shiller, & Ostry, 2014; Pardo, 2006) levels. The Interactive Alignment Model (Pickering & Garrod, 2004) accounts for such effects in speech production and assumes that speech alignment occurs because in order for communication to be successful, mental states of interlocutors should become aligned. If mental states are aligned, interlocutors come to understand the ideas under discussion in the same way. According to the interactive alignment account, alignment percolates between different levels (e.g., phonological, lexical, and syntactic levels) so that alignment on one level stimulates alignment on other levels in both perception and production. Alignment is assumed to be an automatic process in the sense that it is effortless and speakers are unaware of the process. Pickering and Garrod (2004) suggested that alignment comes about through priming of representations between speakers and listeners. In a more recent account of (phonetic) adaptation, Gambi and Pickering (2013) suggested that adaptation occurs because listeners simulate speakers' utterances by constructing forward model predictions of the speakers' utterances using their own speech production system (Pickering & Garrod, 2013). Adaptation to an interlocutor occurs because the listener's predictions mismatch the speaker's utterance and the listener will try to correct the prediction error in perception. Both Pickering and Garrod's interactive alignment model and Gambi and Pickering's simulation theory assume parity between perception and production. Therefore, an adaptation as a consequence of a prediction error in speech perception can lead to adaptations in speech production as well.

Social factors influence the occurrence of phonetic alignment. Babel (2012), for instance, focused on several social variables. Participants first produced a list of target words in a baseline block after which they performed a shadowing task where they repeated words that were presented auditorily over headphones. During the shadowing task, participants either saw a picture of the speaker on the screen or no picture at all. There was more alignment in the social condition (with a picture of the speaker on the screen) than in the auditory exposure only condition. Liking the model speaker (as measured with ratings) also increased alignment. These findings support the view that alignment can be socially driven. However, alignment did not occur to the same extent for each vowel type: There seemed to be more alignment when there was more acoustic space available for alignment. According to Gambi and Pickering (2013), social factors and context factors may influence alignment by affecting how much a listener relies on forward-models of the speaker.

A further important social variable affecting alignment may be the perceived social distance between the interlocutors. One reason for such social distance effects is that comprehension may occur through either a prediction-by-simulation route (simulating interlocutors' speech using one's own production system), or a prediction-by-association route (predicting interlocutors' speech using perceptual experience) (see Pickering & Garrod, 2013 for a detailed discussion). Gambi and Pickering (2013) suggest that in some contexts - for example when an interlocutor is perceived as very different from the listener - listeners may be more inclined to rely on the prediction-by-association route. As this route does not rely on the listener's production system, subsequent speech production is not affected by the predictions made about the interlocutor's speech. This may explain why adjustments in phoneme perception do not always lead to changes in production. For instance, Kraljic, Brennan and Samuel (2008) exposed half of their participants to speech where /s/ was replaced with the pronunciation ~sʃ (ambiguous between /s/ and /ʃ/) when immediately followed by the [tr] (such as in known English dialects). The other participants were exposed to speech in which all instances of /s/ were replaced by ~sʃ (idiolectal condition). There was perceptual learning for the idiolectal variation, but not for the dialectal variation. Importantly, the changes found in perception did not affect subsequent production.

Phonetic Alignment in L2 Speech Production

According to Gambi and Pickering (2013), speech alignment occurs to a larger extent when interlocutors are more similar to each other or when they perceive each other as being more similar. Thus, alignment may fail when interlocutors are highly dissimilar, for example when a non-native speaker is engaged in conversation with a native speaker. Non-native speakers may also lack the flexibility and automaticity in speech production necessary for alignment (Costa et al., 2008), because they may have more limited or erroneous knowledge of L2 linguistic representations and because language perception and production are more effortful in L2.

In line with simulation theory (Gambi & Pickering, 2013), Kim, Horton, and Bradlow (2011) show that closer interlocutor language distance facilitates phonetic alignment. The authors studied alignment in interlocutor pairs with different dialects or with a different L1 with an AXB perceptual similarity test. In this similarity test, an independent group of listeners heard three repetitions of the same target word. The first and last production of the target word represented pronunciation of the target word in the pre- and post-exposure phase (A and B). The second production of the target word (X) was produced by the first speakers' interlocutor. The listeners who judged pronunciation of the target word were asked to decide whether A or B sounded more like X. So, the judgment of the listeners was used as a subjective measure of alignment. Phonetic alignment only occurred when two speakers with the same L1 and dialect were engaged in dialogue and not when the dialects differed or when one conversation partner had a different L1.

Kim, Horton and Bradlow's (2011) finding that alignment was strongest for interlocutor pairs that shared L1 and dialect differs from findings by Hwang, Brennan, and Huffman (2015). These authors studied phonetic alignment in non-native dialogue and asked whether the amount of alignment depended on social affiliation and on the necessity of phoneme disambiguation in dialogue. Unbalanced Korean-English bilinguals interacted with a Korean English-speaking confederate and a monolingual American English-speaking confederate in English. Participants were asked to explain to the confederate how to rearrange a board with words so that it would match that of the participant. Acoustic measures were used to quantify alignment (formant frequencies, closure voicing duration, and vowel duration). Participants produced more English-like phonemes when being immediately primed by a monolingual American confederate pronouncing that same phoneme and their pronunciation did not change when they were speaking to a Korean confederate. Simulation theory can still account for this finding if we assume that the bilingual participants perceived themselves as more similar to the native English confederate than to the Korean confederate. A second experiment showed that participants also produced more English-like phonemes when they needed to distinguish between two potentially ambiguous words on the board.

As in L1, social factors seem to have an influence on the amount of phonetic alignment in L2 speakers. Trofimovich and Kennedy (2014) focused on the nature and the amount of interactive alignment in L2-L2 dialogue. A pair of L2 speakers of English with different L1 backgrounds performed an information exchange task in which interlocutors were required to transmit information unknown to one of the two interlocutors in order to reach a common goal. In line with Kim et al. (2011), alignment was stronger when interlocutors' speech characteristics (fluency, language complexity) were initially more similar. Greater alignment also occurred when interlocutors' affective/personal qualities were initially more similar. This suggests that speakers are perceptive to social context so that similar personality traits lead to an increase in speech alignment (see below).

Kim (2012) observed phonetic alignment of an L1 speaker towards an L2 speaker. In contrast to Kim et al. (2011), who only found alignment in L1-L1 dialogue where speakers shared the same dialect, alignment occurred irrespective of whether the participant shared L1 or dialect with the other speaker. Interestingly, Kim (2012) found that phonetic alignment was larger for larger initial acoustic distances between the two speakers.

Present Study

Most previous studies on phonetic alignment in L2 speakers used subjective measures to test whether interlocutors sounded more alike after an interaction. Here, we will use objective acoustic measures to test whether L2 speakers adjust their speech production of specific phonemes, after being exposed to those phonemes in a sentence context produced by a native confederate. Pickering and Garrod (2013) argue that alignment is a rather automatic process, driven by priming. Hence, an L2 speaker may not only adapt their speech to an L1 speaker, but also vice versa (Kim, 2012). Therefore, we will also test whether a native English speaking confederate aligns her own speech to that of an L2 speaker.

Specifically, we will investigate whether L2 speakers of English adjust their non-native realization of the English phonemes /æ/ and word-final /b/ towards a more native realization after exposure to native realizations of the phoneme. We use word-final /b/ in this study because Dutch non-native speakers of English often replace the English phoneme /b/ with the Dutch phoneme /p/ when it is positioned at the end of a word (Collins & Mees, 1996). This phenomenon exists because Dutch has final devoicing: All voiced consonants in final position are realized as voiceless (Giegerich, 1992). For instance, the English word ‘mob’ /mɒb/ is often mispronounced as /mɒp/. The distinction between the voiced consonants /b d ɡ/ and voiceless consonants /p t k/ in syllable-final position in English is made mainly by vowel length; vowels that precede a word-final voiced consonant are longer than vowels that precede a word-final voiceless consonant (Luce & Charles-Luce, 1985; Raphael, 1972). If alignment occurs, the duration of vowels preceding /b/ should increase with increasing amounts of exposure to native speech. Additionally, closure duration tends to be shorter for voiced word-final stops and longer for voiceless word-final stops (Lisker, 1957; Luce & Charles-Luce, 1985).[2] Therefore, we expect closure duration of /b/ to become shorter when

[2] The duration of voicing in the closure phase of /b/ is often measured to determine voicing of /b/ (e.g., Hwang, Brennan, & Huffman, 2015). However, voicing duration could not be measured reliably due to considerable noise in the recording. Please note that vowel length is the most reliable cue in distinguishing voiced and voiceless final stops (Luce & Charles-Luce, 1985), but for sake of systematicity we also measured closure duration of word-final /b/.

participants are exposed to native speech. Yet, if the confederate aligns with the participant, her vowels preceding /b/ will be shorter whereas closure duration is expected to be longer.

The vowel /æ/ does not exist in Dutch and is often substituted by /ɛ/ by Dutch speakers (Collins & Mees, 1996). To study adaptation in the realization of vowel /æ/, we determine both the first spectral peak (F1) and second spectral peak (F2) as well as the duration of /æ/. F1 correlates with the height of the tongue (vertical tongue position); if the tongue is low (as in /a:/), F1 is high and if the tongue is positioned high (as in /i:/), F1 is low. F2 correlates with the tongue being placed at the front or back of the mouth (horizontal tongue position). In the former placement, F2 is high; in the latter, F2 is low. It is hypothesized that a difference in F1 and F2 before and after exposure should be seen if speech alignment occurs. F1 of /æ/ is slightly higher (lower tongue/jaw position) than F1 of /ɛ/, and F2 of /æ/ is slightly lower (tongue position more back) than F2 of /ɛ/ (tongue position more back). Therefore, if alignment takes place, we expect non-native speakers to adjust their F1 upwards and their F2 downwards when attempting to pronounce the English vowel /æ/. The opposite is expected for the confederate. Also, /æ/ is longer than /ɛ/ (e.g., Bohn & Flege, 1990; Collins & Mees, 1996) and we therefore expect participant to lengthen the vowel if they align with the confederate. However, we expect the confederate to shorten the vowel if she aligns with the participants.

We will also test whether the amount of alignment depends on social context, contrasting a confederate who is present during the experiment with exposure to speech over headphones. The physical presence of the confederate is expected to increase the extent to which participants feel engaged in dialogue, thereby stimulating alignment. Furthermore, we will test whether the amount of phonetic alignment depends on the time lag between perception and production. We expect alignment to be stronger when the time lag between perception and production is short (zero intervening sentences).This would be in line with accounts assuming parity between production and comprehension (Gambi & Pickering, 2013; Pickering & Garrod, 2013) and it would verify Hwang et al.'s (2015) finding that there is alignment when production of a target immediately follows perception.

METHOD

Participants

Thirty-two female students from Ghent University (age M = 25.38, SD = 8.17, range 19 to 57) participated in the experiment in exchange for monetary compensation. They were divided into two groups of 16 (the confederate-absent and confederate-present groups, see below) by random assignment. Because men and women differ in formant frequencies and our confederate was female, we decided to test only female participants. They were all late Dutch-English bilinguals who started learning English around the age of 12 at secondary school for approximately 3-4 hours a week. In addition to this classroom exposure, students in Belgium are regularly exposed to English through television, books, video/computer games, and other kinds of media. All participants were born and raised in Flanders. Proficiency in L1 and L2 was measured using the LexTALE test of vocabulary knowledge for advanced learners of English (Lemhöfer & Broersma, 2012) and a self-report questionnaire. In this questionnaire, participants rated their L1 and L2 proficiency in reading, writing, speaking, and listening on a five-point scale ranging from 1 (*not at all*) to 5 (*perfect/mother tongue*) (see Table 1 for participant characteristics). They also provided more background information on their (previous) place of residence. Besides Dutch and English, all participants also spoke French (mean rating = 3.28 on a scale from 1 to 5 where 1 (*not at all*) to 5 (*perfect/mother tongue*). Participants all reported not to have dyslexia or hearing deficiencies and eyesight was normal or corrected-to-normal.

Table 1

Self-ratings on language proficiency (SD) and LexTALE scores (SD)

Language	listening[a]	Speaking[a]	Reading[a]	Writing[a]	Overall mean[a]	LexTALE
Dutch						
Confederate present	5 (0.00)	5 (0.00)	5 (0.00)	5 (0.00)	5 (0)	92.11 (4.49)
Confederate absent	5 (0.00)	5 (0.00)	5 (0.00)	5 (0.00)	5 (0)	85.16 (14.65)
English						
Confederate present	2.56 (.50)	2.38 (.77)	2.69 (.42)	2.31 (.77)	2.48 (.80)	76.80 (12.62)
Confederate absent	2.69 (.60)	2.31 (.60)	2.50 (.63)	2.25 (.58)	2.44 (.54)	70.94 (12.49)

Note. There were no significant differences between English proficiency scores in the confederate absent and confederate present groups (all p-values >.1). The difference between the proficiency scores for Dutch and English was significant in each condition (all p-values <.0001).

[a] Ratings were given on a scale from 1 to 5 with 1=not at all and 5= native speaker.

Confederate

The confederate was female and she originated from the Pacific Northwest of the United States of America. She was 30 years old at the time of testing and had been living in Belgium for little over a year. English was her native language but she also spoke French and Dutch. The confederate also performed the LexTALE in Dutch and in English. Her score for Dutch was 67.5 and her score for English was 96.25.

Design

The experiment consisted of three blocks: a baseline block, an exposure block, and an alternating block. In the baseline block, 30 sentences, each with two target words (one for /æ/ and one for /b/) were presented to the participant to read out loud. In the exposure block 30 different sentences with the same 60 target words were read out loud by the confederate. In the post-exposure (alternating) block, the participant and the confederate alternated in reading 120 sentences out loud that each contained one of the 60 target words. Over the course of the experiment, each target word occurred four times (produced twice by the confederate and twice by the participant) but it was presented in a different sentence each time.

In the alternating block, the lag between the sentence containing a target word that was produced by the confederate and the test sentence containing that same target word produced by the participant could be either zero or four. A lag of zero means that the critical sentence for the participant was presented immediately after the confederate produced a sentence containing the same target word. A lag of four indicates that four intervening sentences were presented between the critical sentences of the participant and confederate. Lag was a within-participant variable. To enable the lag manipulation, 30 fillers were added to the 120 sentences in the alternating block. These filler sentences had a similar structure and length as the critical sentences but they did not contain the target words or the specific contrast. Half of the fillers were read by the participant and half of the fillers were read by the confederate. Each phoneme was presented fifteen times at lag zero and fifteen times at lag four in the alternating block.

There was a condition in which the confederate was present in the same room as the participant during the experiment, and a condition in which the confederate was not present in the same room but read out loud sentences in a microphone (Røde USB 1000A) in another room (see Procedure for details). This social context (confederate present or absent) was manipulated between participants. Table 2 below summarizes the design of the experiment.

Table 2

Design of the experiment

Block	sentences	Speaker	Lag	Social context
Baseline block	30 sentences with words including /æ/ and word-final /b/	Participant	No lag	Confederate present/absent
Exposure block	30 further sentences with same targets as baseline block	Confederate	No lag	
Alternating block	60 further sentences (targets appeared twice in this block: once for participant and once for confederate) + 30 fillers	Participant + Confederate	Lag 0 + lag 4	

Materials

There were two target phonemes: word–final /b/ and the vowel /æ/ (see Appendix 5A for the full stimuli list). We selected 30 English target words for each of the two phonemes. English /æ/ (as in 'map' and 'trap') is affected by dark [ɫ], giving a retracted [ä] such as in pal, shall. The mouth is not as open when pronouncing English /æ/ before velar phonemes /ŋ, k, ɡ, ʍ, w/ giving rise to [æ̝] (e.g., back, bag, bang) (Collins & Mees, 1996). Therefore, the vowel was never followed by one of these sounds in a target word. In addition, /æ/ was never word-initial. /b/ was always preceded by a vowel in a target word (as in 'tub' and 'job').[3] The target words never occurred at the end of a sentence, or before /f/ and /v/ because the /b/ becomes a labial-dental sound if it precedes these phonemes (as in 'obvious') (Collins & Mees, 1996). Therefore, /b/ was always followed by a vowel.

Each /b/-target word was randomly paired with an/æ/-target word in a sentence for the baseline blocks. In the exposure block, the /b/-target word was again randomly paired with another /æ/-target word in another sentence, resulting in 60 sentences containing one instance of each contrast created for the first two blocks. An additional two sentences containing only one target word were created for each target word for the post-exposure block. There were no particular constraints on the sentences: They were constructed by the authors, both long and short sentences were included, and the sentences were non-constraining towards the target words. The confederate checked whether the sentences were grammatically correct before the experiments were run; she corrected one sentence.

Two presentation lists were created for each block where the sentences were presented in pseudorandom order: The pattern of the lag manipulation in the alternating block was the same for both lists, but the order of the sentences was randomized. Each list could be presented in version A or B so that the sentences read by the participant in version A were read by the confederate in version B and vice versa.

Procedure

In the confederate-present context, the experimenter went to pick up the participant and the confederate in the hall of a university building. Throughout the experiment, the confederate acted as if she was just another participant and the confederate did not speak English before the experiment started. In the confederate-absent context, the confederate was seated in another room and the participant did not see the confederate during the experiment. In this condition, participants were told that they would be listening to recordings of spontaneous speech and participants thought they were the only one being tested. Participants received oral and written instructions in Dutch to read aloud the English sentences presented on the screen. We told the participants that the experiment tested whether comprehension of sentences was better when participants read the

[3] For one sentence with a target word that ended in word-final /b/, the word 'while' followed the target word ('stub') instead of a word starting with a vowel. As the intercept 'sentence' was included in the linear mixed effects model this should not lead to problems in the analyses.

sentences or when someone else read the sentences (i.e., the confederate). This explanation was provided to draw the participants' attention away from the true goal of the experiment.

Participants were tested in a silent room and were seated in front of a computer screen and a microphone while wearing headphones. In both the confederate-present and the confederate-absent context, the participant, the confederate, and the experimenter each worked on a laptop computer. The experimenter used his laptop to record the speech of the confederate and the participant. The confederate's and participant's laptop were used for visual stimulus presentation by means of the computer program E-prime 2.0. The confederate's and the participant's microphones were connected to a mixer, which was in turn connected to the experimenter's laptop. The recordings were made in *Audacity* with a sampling frequency of 48 kHz. The participant and confederate heard each other live over headphones both when the confederate was present in the same room and when she was seated in the other room. None of the participants in the confederate absent context noticed that the confederate's speech production was live instead of a recording. The confederate's speech was live in both conditions to keep the conditions as similar as possible on all variables except for physical confederate presence; pronunciation of the sentences was of comparable variability and the confederate could also hear the participant's speech in both versions.

Table 2 summarizes the design. In the baseline block, participants read the sentences out loud, while the sentences were read by the confederate in the exposure block. In the alternating block, the participant and confederate each read a sentence in turn. Every trial started with a fixation cross on the screen, after which a sentence was presented if it was the participant's turn to read a sentence. When the confederate read aloud a sentence, a picture of an ear and the text 'Listen' was presented on the participant's screen. The sentence or the word 'Listen' remained on the screen until the participant pressed a button, after which the next sentence was presented. A comprehension question was presented after 10% of the sentences. The participant and confederate (when present) were asked to answer the questions by pressing the F-button if the statement about the sentence was incorrect and the J-button if it was correct. To ensure that the participant and the confederate continued at the same pace with the next trial, they were asked to say 'okay' before continuing after answering a question. Only the participant was asked to say 'okay' after answering a question in the confederate absent context. After the experiment, participants were asked whether they thought they knew what the experiment was about. None of the participants suspected that the experiment was about their pronunciation, and hence neither about alignment.

Acoustic Measures and Annotation

Analyses were performed on the recordings of the participants' speech. The target sounds were annotated by hand using Praat (Broersma & Weenink, 2014) after which a script was used to extract the formant frequencies of the first and second spectral peaks (F1 and F2) and the length of annotated vowel and word segments. For /æ/, both the vowel itself and the entire word were annotated. For word-final /b/, the preceding vowel, closure duration, and the entire word were annotated.

Phoneme boundaries were determined as accurately as possible through visual and auditory inspection. Vowel boundaries were placed at F2 onset and offset in the spectrogram or, if F2 onset or offset was unclear, where two or more formants appear or drop out together (Hwang et al., 2015). The offset of the target word with /b/ was always set right after the release burst of /b/. If the release was not audible and/or visible, it was placed immediately before the onset of the next word. Closure duration was defined as the length of the segment from vowel offset until the release burst. If the release was not visible and/or audible, closure duration was not taken into account.

The Praat script determined the formants using a 0.00625s time step and a 0.025s window length. Formant frequencies were then aggregated so that the dataset contained one mean formant frequency for F1 and F2 for each produced phoneme (see Appendix 5B for a table displaying raw values of formant frequencies and durations). To be able to create a measure of /æ/ that was normalized to each participant's vowel space, we also annotated all occurrences of /ɛ/ in the experiment. Depending on the list, there were 17 or 22 occurrences of /ɛ/ in the baseline block and 43 or 52 occurrences of /ɛ/ in the post-exposure block. The frequencies of F1 and F2 of /æ/ and /ɛ/ were transformed to the psychoacoustic Bark scale for analysis (Traunmüller, 1990). The participants' F1 and F2 values of /æ/ were then divided by the mean F1 and F2 formant frequency of /ɛ/ (respectively) in the same block to create the normalized measure. This measure is more informative than plain F1 and F2 of /æ/, because it shows to what extent L2 speakers of English distinguish between /æ/ and /ɛ/. The experimental set-up induced considerable noise to the recordings. However, the spectrograms showed that the speech signal was considerably stronger than the noise signal.

The duration measures used for the analyses of the production of the vowel preceding word-final /b/, closure duration, and /æ/-duration were relative (the duration of the vowel/closure divided by the duration of the word). This relative measure of vowel length was used to correct for speech rate. In the analyses, when we refer to F1, F2 or duration, we always refer to the normalized measures. All values above and below 2.5 standard deviations of a participant's mean for an item were excluded from the analysis.

Annotation took approximately 250 hours; the task was divided over five researchers. Interclass correlation (ICC) was calculated for all duration measures based on the pre-exposure block of a randomly selected subject by means of the package ICC in R (3.4.1) (R Core Team, 2013). ICC was only calculated for duration because segment duration directly reflects placement of phoneme boundaries. Two-way models were used with type 'agreement' and unit as definition. There was a high degree of reliability between phoneme boundary placement for almost all measures (see Table 3 below).

Table 3

Interclass correlation information on five different measures. F- and p-values indicate whether the correlation significantly differs from zero.

Measure	ICC	Lower CI (2.5%)	Upper CI (97.5%)	F-value (df)	P-value
Word duration /æ/	.821	.678	.913	33.6 (22, 33.1)	< .001
Vowel duration /æ/	.672	.499	.823	12.8 (21, 68.9)	< .001
Word duration /b/	.825	.699	.910	31.5 (24, 45.8)	< .001
Vowel duration before /b/	.823	.700	.902	22.7 (23, 95.9)	< .001
Closure duration	.209	-.055	.703	2.34 (5, 23.3)	.074

The ICC of closure duration is low because of many missing values in the measurements (where only five instances of closure duration were measured by one of the annotators). The release of the /b/ was not always audible and/or visible and therefore this particular measure has more missing data. The percentage of annotated closure durations amounted to 69.5% (1335/1920) in the confederate data set and to 60.9% (1169/1920) in the participant data set.

Analyses

We first determined whether there were substantial differences between the Participant's and the Confederate's[4] acoustic characteristics for each target phoneme. Then, we tested whether Participants' phonetic characteristics changed after exposure to the Confederate's speech by comparing the post-exposure (alternating) block and the pre-exposure (baseline) block, and whether the degree of change depended on social context (the presence or absence of a Confederate during the experimental session). Additionally, we tested whether phoneme production in the post-exposure block was more similar to that of the Confederate immediately after the Participant had heard the Confederate's production of the phoneme (lag 0) than when four sentences intervened between perception and production (lag 4). Finally, we tested whether mere repetition of the target sounds lead to changes in Participants' production by assessing change over the course of the baseline block and whether listening to and producing target phonemes in the post-exposure block lead to additional changes over the course of that block (trial number effects).

We ran the same analyses for the Confederate and additionally tested whether she also changed her target phoneme production over the course of experimental sessions (one Participant was tested each experimental session). For the Confederate, trial number effects were only assessed in the post-exposure block in order to test whether more interaction with the participants led to (more) adaptation over the course of the post-exposure (alternating block). Whether mere repetition of the target phoneme lead to changes in the confederate's target phoneme productions was not of interest here. Additionally, the exposure block was not a true baseline block like the baseline block for the participants (because the confederate already heard the participant's production during the baseline block at this point). Therefore, we did not assess the effects of trial number in the exposure block.

Our data set was analyzed with linear mixed effects models in R (version 3.4.0). P-values for the fixed effects and interactions in the final models were computed using the lmerTest package (version 2.0-33) (Satterthwaite degrees of freedom approximation) (Kuznetsova, Brockhoff, & Christensen, 2016). First, we ran a simple model for the normalized measures of duration of the vowel preceding word-final /b/, closure duration F1, F2, and /æ/ duration separately. These simple models included the three main experimental fixed factors social context (confederate absent/present), block (pre-exposure and post-exposure), and list (control variable: the different stimuli presentation lists) as well as the interaction between social context and block. The random intercepts were participant, word, and sentence. Participants' L2 proficiency (centered LexTALE

[4] In the Analyses and Results section we use Participant and Confederate (with capital letter) to refer to experimental role. The terms are not capitalized when they refer to experimental factors (by participants random intercept or confederate absent/present condition).

score) and the two- and three-way interactions between proficiency, block, and context (participant data only) were only added to the model if they contributed to the model fit. Similarly, experimental session and the interactions between session, block, and context were only added to the models for the confederate data set if they contributed to the model fit. Note here that every session had a new Participant but the same Confederate. Participant was never included as random intercept when session was a fixed factor in the model because the intercept captured the same information.

Subsequently, random slopes were determined by comparing models with and without each random slope with a Chi-square test (Baayen, 2008). If the models differed significantly, then the model that explained the most variance and with the lowest AIC value was used. Random slopes were tested in a fixed order (block, condition, list, trial number if applicable). Also, the random effects structure was simplified if running the model resulted in convergence errors.

A separate linear mixed effects model was constructed for the data from the post-exposure block to test whether there was an effect of Lag. This model included the fixed factors Lag (0 or 4 sentences), social context (confederate present or absent), and their interaction. The control variable presentation list was also included as a fixed factor. Random intercepts of participant, word, and sentence were included and random slopes were once again determined by model comparison. The effects of trial number were also assessed separately in the baseline and post-exposure block. The data sets and scripts used for the analyses are available online at Open Science Framework (https://osf.io/p62j4/).

RESULTS

Target Phoneme /b/

Figure 1 below shows that the Confederate produced longer vowels preceding word-final /b/ than the Participants, whereas Figure 2 shows that the Participants had a longer mean closure duration than the Confederate. Two linear mixed effects models with speaker (Confederate vs. Participant) as fixed factor, session as random intercept, and random slope for session were constructed to test whether these differences were significant. The differences were significant for both vowel length (β = -.11, SE = .008, t = -13.29, p <.001) and closure duration (β = .09, SE = .04, t = 20.14, p <.001).

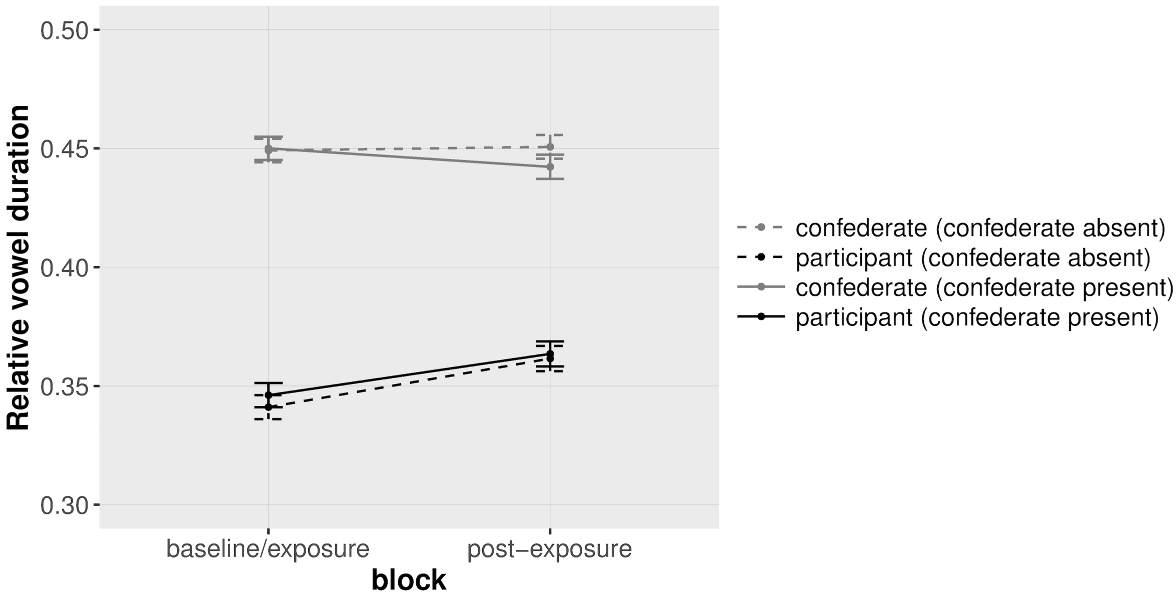


Figure 1. Relative vowel duration of the vowel preceding word-final /b/ in the baseline and post-exposure block for the Participant and in the exposure and post-exposure block for the Confederate. Error bars denote standard errors.

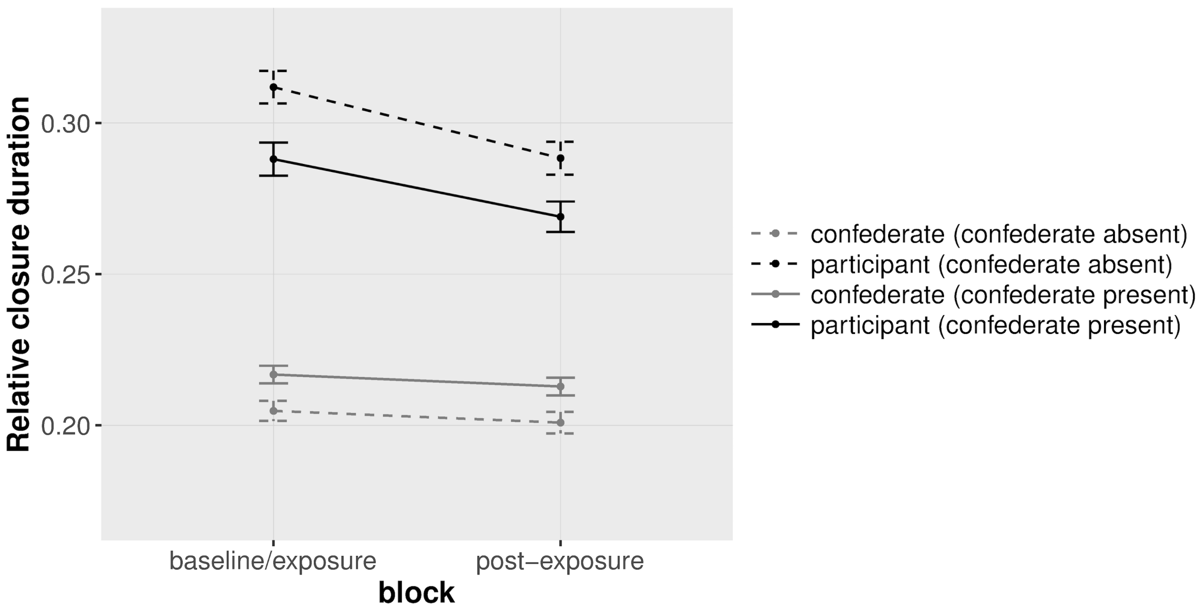


Figure 2. Relative closure duration of the vowel preceding word-final /b/ in the baseline and post-exposure block for the Participant and in the exposure and post-exposure block for the Confederate. Error bars denote standard errors.

Participants. There was a main effect of block (baseline vs. post-exposure) for the Participants on both measures (preceding vowel duration: (β = .02, *SE* = .004, *t* = 4.56, *p* < .001); closure duration: β = -.02, *SE* = .005, *t* = -3.59, *p* < .001). The duration of the vowel preceding word-final /b/ increased after exposure to the Confederate's speech, and the closure duration of the Participants decreased. Thus, the Participants' production of final /b/ became more like the Confederate's production on both acoustic measures. The interaction between block and social context was not significant (preceding vowel duration: β = -.002, *SE* = .006, *t* = -.35, *p* = .72; closure duration: β = .004, *SE* = .007, *t* = .57, *p* = .57) (full results are presented in Table C1 in Appendix 5C). Finally, L2 proficiency did not improve the model fit (preceding vowel duration: $\chi^2(4) = 8.1$. *p* = .09; closure duration: $\chi^2(4) = 7.05$, *p* = .13).

There was no main effect of trial number in the baseline block (*p-values*>.1). The post-exposure block, however, did reveal a main effect of trial number on vowel duration only (β = .001, *SE* = .0004, *t* = 2.45, *p* = .01); the vowel preceding /b/ became longer over the course of the post-exposure (alternating) block. There were no interaction effect between trial number and social context in either the baseline or post-exposure block (*p-values*>.05).

The main effect of lag did not reach significance (preceding vowel duration: β = .009, *SE* = .007, *t* = 1.26, *p* = .21; closure duration: β = .001, *SE* = .008, *t* = .16, *p* = .88), nor did the interaction of lag and social context (preceding vowel duration: β = -.006, *SE* = .009, *t* =-.7. *p* = .48; closure duration: β = -.0005, *SE* = .009, *t* = -.05, *p* = .96).

Confederate. Experimental session improved the model fit for preceding vowel duration ($\chi^2(4) = 20.38$, *p* < .001) and for closure duration ($\chi^2(4) = 9.89$, *p* = .04). This factor was therefore included in the final models. There was only a significant main effect of session for vowel duration (β = -.0009, *SE* = .0003, *t* = -2.95, *p* = .003), with the Confederate's relative vowel length decreasing over the course of experimental sessions. No other main effects or two- and three-way interactions between session, block and social context were significant for closure duration or preceding vowel duration (all *p-values*>.05) (full results are presented in Table C2 in Appendix 5C). The main effect of trial number did not reach significance in the post-exposure block and there was no interaction between trial number and social context on either measure (*p-values*>.1).

Summary target phoneme /b/. Participants showed an adaptation effect for both the duration of the vowel preceding word-final /b/ and closure duration. The increase in the Participants' duration of the vowel preceding word-final /b/ over the course of the post-exposure block suggests that Participants adapted vowel length more after hearing and producing more target sounds. No effects of social context or lag were found. The Confederate did not adapt the duration of these measures to the Participants' productions from the exposure block to the post-exposure (alternating) block, but she did significantly shorten her vowels preceding /b/ (they became closer to the Participants' vowel length) after taking part in more experimental sessions.

Target Phoneme /æ/

Figure 3 to 5 show the normalized mean F1 scores, F2 scores, and the relative duration of /æ/ for the Participants and the Confederate in the confederate-present and -absent contexts before and after exposure. Figure 3 shows that, as expected, the Participants' mean F1 was lower than the Confederate's. A linear mixed effects model with speaker (Confederate vs. Participant) as a fixed factor and session as random intercept that was run for the baseline and exposure block data confirmed this (β = -.05, *SE* = .005, *t* = -10.43, *p* <.001). The Participants' F2 values in the baseline block were also significantly different from the Confederate's F2 values in the exposure block (β = -.008, *SE* = .004, *t* = -2.20, *p* = .036) (Figure 4). A final model indicated a significant difference in mean duration of /æ/ between Participants and Confederate (β = -.07, *SE* = .009, *t* = -7.32, *p* <.001) (Figure 5).

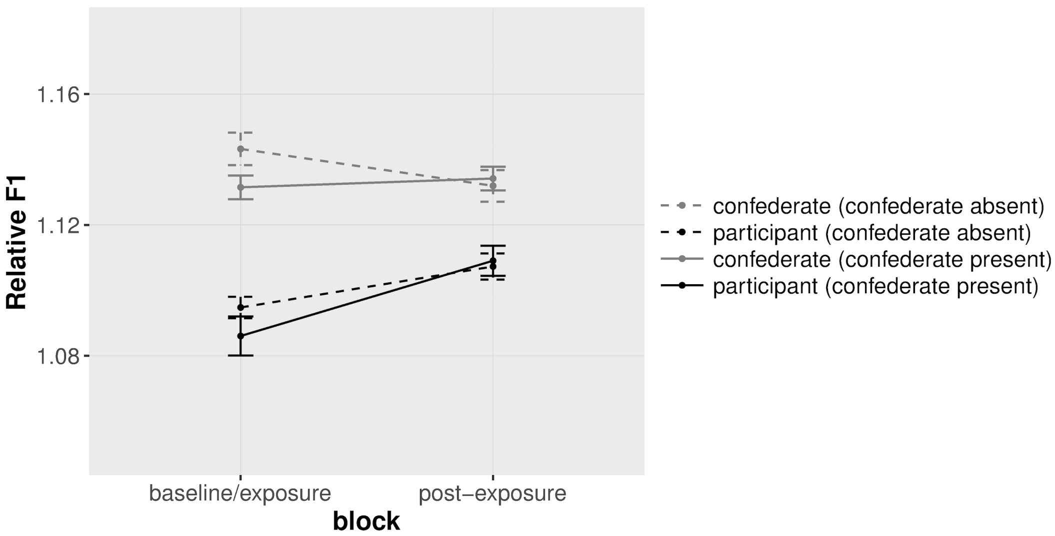


Figure 3. Relative F1 frequencies of target vowel /æ/ in Bark in the baseline and post-exposure block for the Participant and in the exposure and post-exposure block for the Confederate. Error bars denote standard errors.

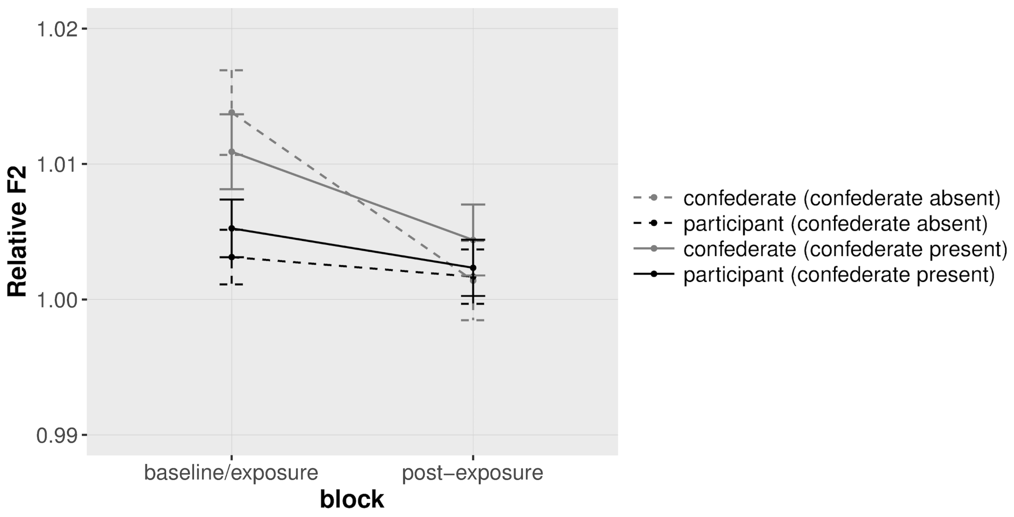


Figure 4. F2 frequencies of target vowel /æ/ in Bark in the baseline and post-exposure block for the Participant and in the exposure and post-exposure block for the Confederate. Error bars denote standard errors

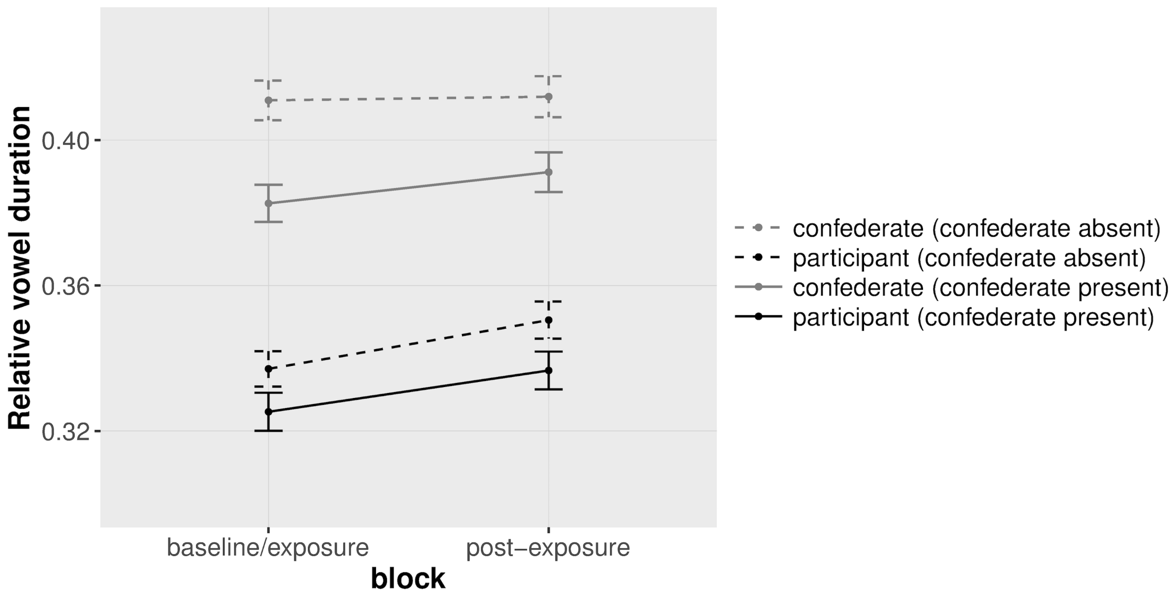


Figure 5. Relative vowel duration of the target vowel /æ/ (duration of the vowel divided by duration of the word) in the baseline and post-exposure block for the Participant and in the exposure and post-exposure block for the Confederate. Error bars denote standard errors.

Participants. The difference between the Participants' production of /æ/ in the baseline and post-exposure block was not significant for any of the acoustic measures (F1: β = .01, *SE* = .01, *t* = 1.36, *p* = .18, F2 β = -.002, *SE* = .002, *t* = -.67, *p* = .16; /æ/ duration: β = .01, *SE* = .008, *t* = 1.62, *p* = .11). The interaction between block (baseline vs. post-exposure block) and social context (confederate present vs. confederate absent) was not significant either (F1: β = .01, *SE* = .01, *t* =

.84, p = .41; F2: β = -.001, SE = .003, t = -.44, p = .66; /æ/ duration: β = -.0002, SE = .01, t = -.03, p = .98). Full results are presented in Table C3 of Appendix 5C.

L2 proficiency and the interactions between proficiency, block, and social context did not improve the model fit for any of the acoustic measures (F1: $\chi^2(4) = 3.03$, p = .55; F2: $\chi^2(4) = 1.69$, p = .79, duration: $\chi^2(4) = 4.28$, p = .37). The main effect of trial number was not significant in the baseline block, nor in the post-exposure (alternating) block for any measure (*p-values*>.1). The interaction between trial number and social context was not significant either (*p-values*>.05).

There was no effect of time lag between perception and production (zero vs. four intervening sentences) in the post-exposure block (F1: β = -.001, SE = .008, t = -.14, p = .89; F2: β = .002, SE = .003, t = .46, p = .65; duration: β = -.003, SE = .008, t = -.36, p = .72) , or an interaction between time lag and social context (F1: β = .006, SE = .01, t = .6. p = .55; F2: β = -.007, SE = .004, t = -1.66, p = .10; duration: β = -.004, SE = .008, t = -.53, p = .60).

Confederate. As for the Confederate, there was no significant effect of block on F1 (β = .009, SE = .008, t = 1.09, p = .27) or on duration (β = .003, SE = .006, t = .43, p = .67). The Confederate did significantly decrease her F2 from the exposure block to the post-exposure (alternating) block (β = -.016, SE = .005, t = -2.95, p = .004). A main effect of social context was found for F2 (β = -.02, SE = .004, t = -6.15, p < .001), with the Confederate's F2 being lower in the present than in the absent condition. Social context was also significant for duration (β = -.03, SE = .01, t = -2.48, p = .02), with the Confederate producing shorter vowels in the present condition. The interaction between block and social context was not significant (F1: β = -.008, SE = .009, t = -.89, p = .37; F2: β = -.008, SE = .005, t = 1.52, p = .13; duration: β = .005, SE = .006, t = .9. p = .37.

The factor experimental session and the two- and three-way interactions of session, block and social context improved the model fit for F1 ($\chi^2(4) = 26.54$, p < .001) and for F2 ($\chi^2(4) = 77.55$, p <.001) and were therefore included in the final models for those measures. The effect of session did not contribute to the model fit for duration ($\chi^2(4) = 4.14$, p = .39). There was a three-way interaction between block, social context, and session for F1 (β = .001, SE = .0004, t = 2.68, p = .008). Post-hoc tests with lsmeans showed that in the exposure block in the absent condition, F1 increased significantly over the course of experimental sessions (β = .0007, SE = .0002, t = 2.87, p = .008), but not in the post-exposure block (β = -.0004, SE = .0002, t = -1.82, p = .13). In the present condition there was a significant decrease of F1 over sessions both in the exposure (β = -.0006, SE = .0002, t = -2.96, p = .006) and the post-exposure block (β = -.0005, SE = .0002, t = -2.52, p = .02). There was a main effect of session on F2 (β = -.0008, SE = .0001, t = -5.62. p <.001) and session also interacted with condition (β = .001, *SE* = .0002, *t* = 6.49, *p* <.001). Post-hoc tests with lsmeans revealed a positive trend for session in the present condition and a negative trend in the absent condition. Full results are presented in Table C4 in Appendix 5C.

The effect of trial number in the post-exposure block was not significant for the F1 (β = -.0007, *SE* = .0004, *t* = -1.74, *p* = .09), but there was a main effect of trial number for the F2 in the post-exposure block β = -.0006, *SE* = .0002, *t* = -3.51, *p* < .001). This suggests a further downward change of F2 over the course of the post-exposure (alternating) block. The effect of trial number was also significant for vowel duration in the post-exposure block (β = -.001, *SE* = .0004, *t* = -3.48, *p* <.001): the Confederate shortened her vowels over the course of the post-exposure block. There were also a significant interaction between condition and trial number for F2 (β = .004, *SE* = .0002, *t* = 2.10, *p* =.037) and for vowel duration (β = .001, *SE* = .0004, *t* = -2.36, *p* <.02), indicating that the adjustment over trials was larger in the absent than in the present condition.

Summary target phoneme /æ/. Participants did not show a change in their pronunciation of /æ/ after exposure to /æ/ pronounced by the Confederate. Time lag between perception and production of /æ/ did not affect pronunciation either. The confederate lowered her F2 from exposure to post-exposure, but there was no change in her F1, or vowel duration. The Confederate's F2 was lower and her vowel duration shorter in the present than in the absent condition. Further, the confederate's F1 increased over sessions in the exposure block in the absent condition, and decreased over sessions in the present condition. The confederate's F2 increased over sessions in the present condition and decreased over sessions in the absent condition.

DISCUSSION

Aligning with a native speaker may be a useful mechanism for language learning. On the other hand, L2 speakers may be too dissimilar from native speakers for phonetic alignment to occur. The aim of the present study was to test whether unbalanced Dutch-English bilinguals adapt their L2 speech after listening to a native speaker of the target language. Additionally, we tested whether a native English confederate also adapted her pronunciation to our (non-native) participants' pronunciation. In particular, we focused on the pronunciation of the phoneme /æ/ and the vowel preceding word-final /b/ in English.

There was significant alignment of the participant to the confederate for closure duration of word-final /b/ and duration of the vowels preceding word-final /b/. Specifically, closure duration of the participants was shortened in the post-exposure block compared to baseline whereas the duration of the participants' vowel preceding word-final /b/ increased. However, there was no alignment for the other target phoneme, /æ/. No main effect of block was seen on the F1, F2, or duration of /ae/ for the participant. Social context did not affect alignment of either phoneme, nor did time lag between perception and production.

The finding that L2 speakers of English adapt their pronunciation of word-final /b/ and the preceding vowel supports the findings of Hwang et al. (2015), who also found alignment of L2 speakers in L2-L1 dialogue for /b/ (on preceding vowel duration but not closure voicing duration) and /æ/ (on vowel duration and F1 but not F2). It also strengthens the claim that alignment takes place when speakers can improve their L2 pronunciation by adapting to L1 speech. As demonstrated by the lack of a trial number effect in the baseline block, the adaptation of word-final /b/ was not merely an effect of repeated production of the phoneme.

However, the lack of alignment on the target vowel /æ/ suggests that alignment by L2 speakers does not occur under all circumstances. Perhaps our participants could not sufficiently perceive the difference between their own speech and that of the native speaker. Dutch native speakers often have difficulty distinguishing /æ/ and /ɛ/ in speech perception (Broersma, 2005; Weber & Cutler, 2004). If the difference in pronunciation cannot always be perceived by Dutch speakers, then it might be very hard if not impossible for them to adjust their phoneme boundaries of this particular vowel. In contrast, Dutch listeners have no difficulty distinguishing /b/ and /p/, as /b/ does occur in Dutch (only not at the end of the word).

The acoustic characteristics of the participants' word-final /b/ in the post-exposure block were not affected by the number of sentences (zero or four) intervening between the participants' and confederate's production of the target phoneme. This finding extends the observations of Hwang et al. (2015), who found alignment in L2 speech after immediate priming by the L1 confederate without including a lag between target words. We found alignment of word-final /b/ both in the immediate condition (lag 0) and the delayed condition (lag 4). An account in terms of automatic priming would predict time lag effects. Possibly, the influence of an exposure to a native phoneme is relatively long-lasting, so that the confederate's production four trials back still affects the participant's current production. However, it is also possible that the cumulative influence of the confederate's productions during the exposure phase was strong enough to last during the post-exposure phase, so that any new exposure (whether immediate or delayed) had little further effect. Also, simulation theory (Gambi & Pickering, 2013) would predict that when episodes of comprehension are tightly interwoven with episodes of production (like in our post-exposure/alternating block), simulation should be enhanced. This would perhaps not predict an effect of time delay between perception and production of a specific phoneme, but an effect of time delay between speech perception and production in general. In the post-exposure block in our study, the time delay between speech perception and production was always short. The effect of trial number on the length of the vowel preceding word-final /b/ (further lengthening of the vowel over the course of the post-exposure block), supports this claim.

The present study also tested whether there was a difference in the amount of phonetic alignment between an L2 speaker and an L1 speaker when the L1 speaker was physically present or absent. Based on the Interactive Alignment Model, priming should result in alignment, irrespective of the social context. However, if alignment is not solely based on priming but is also modulated by contextual factors (e.g., social context, motivation, or beliefs about an interlocutor), the presence of a confederate may boost alignment. Hence, we hypothesized that the actual *presence* of the confederate would have an influence on the amount of phonetic alignment. Yet, no social context effects (effects of confederate presence) were found for the participants. Gambi and Pickering (2013) suggest that phonetic adaptation through simulation depends on the allocation of limited attentional resources. Perhaps in our study, due to disadvantages associated with L2 processing, the nonnative speakers had less resources available to further adjust their pronunciation of /b/ to that of the confederate's when the confederate was present.

The confederate did not show consistent alignment with the participants.[5] The confederate adjusted her F2 of /æ/ from the exposure block to the post-exposure (alternating) block (in the direction of the participant mean). Within the post-exposure block (the alternating block) the confederate further lowered her F2 value and she also shortened the vowel /æ/ over the course of the post-exposure block (in the direction of the participants). The confederate's pronunciation also changed over experimental sessions, but there was no systematic convergence with the participants. The confederate's vowel before /b/ became shorter across sessions (closer to the participants). Her F1 of /æ/ increased across sessions in the absent condition in the exposure block (diverging from

[4] There were five participants whose mean relative F1 was higher than the confederate's at baseline, and there were 12 participants whose mean relative F2 was higher than the confederate's at baseline. We conducted additional analyses where participants with a higher mean F1 and F2 value at baseline were excluded. As their initial F1 and F2 values were higher than that of the confederate, one would not expect to see phonetic alignment in these participants (or maybe even reversed alignment). However, no main effects of block or interactions between block and social context were found (*all p-values*>.05).

the participants), but not in the post-exposure block. She also lowered her F1 across sessions in the confederate present condition in both blocks (converging with the participants). F2 became significantly lower across sessions in the absent condition (diverging from the participants) and higher in the present condition.

Our findings partially support the Interactive Alignment account (Pickering & Garrod, 2004) which assumes that alignment is a rather automatic process. On the one hand, it is supported specifically by the findings that there was no support for a modulation of alignment on /b/ by social context, suggesting that alignment occurs automatically without considering the situation. Moreover, the participants were unaware of the goal of the experiment. On the other hand, the finding that the confederate did not align her speech towards that of the participants does not support the automaticity of alignment.

Simulation theory (Gambi & Pickering, 2013) can account for this apparent inconsistency if we assume that L2 speakers aspire to be more similar to L1 speakers (and therefore perceive themselves as being more similar), whereas L1 speakers perceive themselves to be very dissimilar from L2 speakers. Gambi and Pickering (2013) suggest that when the perceived difference between two interlocutors is too large, interlocutors may rely less on simulations of the other person's speech. If less simulation occurs during speech comprehension, then there should also be less influence of simulations on one's own speech production. Also, simulation of an L2 speaker's utterances by an L1 speaker may simply fail because the L1 speaker lacks experience with the L2 speaker's utterances. Even though word final /p/ exists in English like in Dutch (e.g., *hip hop)*, the devoiced pronunciation of word final /b/ in English words (e.g., *blop* instead of *blob)* by L2 speakers may be unfamiliar/unexpected to a native speaker. Therefore, the L1 speaker may be slow to adjust her predictions of the L2 speakers' utterances and therefore alignment may fail.

The current study focused on alignment by L2 participants, rather than native speakers, and therefore only included one confederate. Therefore, the lack of consistent alignment in the confederate's speech data set might also be due to individual characteristics of the confederate. In this study we wanted the participants to be exposed to the same speaker to reduce variability, but future research with multiple confederates could point out whether adaptation of an L1 speaker towards an L2 speaker occurs under some circumstances.

Gambi and Pickering (2013) suggest that when there is more information available at linguistic levels other than the phoneme level, limited availability of attentional resources may cause predictions to be based on those levels (such as the word or sentence level). Phonetic imitation may therefore be less pronounced. Perhaps this can explain why Hwang et al. (2015) did find adaptation of /æ/ in non-native speakers when primed by a native speaker, whereas we did not. In their experiment, confederate utterances were very simple (e.g., "what is below Hob"). Our stimuli contained longer and more complex sentences and participants may therefore have made use of predictions at other linguistic levels, making them less sensitive to variations at the phonetic level.

A potential limitation of the current study is that the baseline block was not entirely identical across conditions. During the baseline blocks, the confederate was present in the same room as the participant in the confederate present condition whereas, she was absent in the absent condition. The sole presence of the confederate might have influenced pronunciation of the participant in the baseline block, for example by motivating the participant to produce the sentences with a more native-like accent. That being said, the confederate did not speak English (nor Dutch) up until the exposure block, meaning that the confederate's speech could not have affected the participants' utterances at baseline. Moreover, there was no main effect of social context nor an interaction between block and social context for the measures that showed alignment (vowel duration preceding word-final /b/ and closure duration). We therefore argue that this inconsistency would not have greatly affected the results.

In conclusion, results from the current study show that speech production in L2 is influenced by exposure to speech produced by a native speaker of that language. However, the effect depended on the particular phoneme, possibly related to the degree to which participants can perceive the relevant phonemic distinction. Adaptations seem to last over at least four intervening trials. There was no compelling evidence that such influences are affected by social factors.

REFERENCES

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge University Press.

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, *40*(1), 177–189.

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2012). Effects of phonological feedback on the selection of syntax: Evidence from between-language syntactic priming. *Bilingualism: Language and Cognition*, *15*(03), 503–516.

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2013). From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition*, *127*(3), 287–306.

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, *11*(03), 303. https://doi.org/10.1017/S0142716400008912

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*(2), 707–729. https://doi.org/10.1016/j.cognition.2007.04.005

Branigan, H. P., Pickering, M. J., Pearson, J., McLean, J. F., & Brown, A. (2011). The role of beliefs in lexical alignment: Evidence from dialogs with humans and computers. *Cognition*, *121*(1), 41–57. https://doi.org/10.1016/j.cognition.2011.05.011

Broersma, M. E. (2005). *Phonetic and lexical processing in a second language*. Radboud Universiteit Nijmegen.

Broersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer (Version 5.3.78). Retrieved from www.praat.org

Collins, B., & Mees, I. (1996). *The phonetics of English and Dutch* (3. rev. ed). Leiden [u.a.]: Brill.

Costa, A., Pickering, M. J., & Sorace, A. (2008). Alignment in second language dialogue. *Language and Cognitive Processes*, *23*(4), 528–556.

Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time. *The Journal of the Acoustical Society of America*, *119*(4), 1950. https://doi.org/10.1121/1.2178721

Gambi, C., & Pickering, M. J. (2013). Prediction and imitation in speech. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00340

Giegerich, H. J. (1992). *English phonology: an introduction*. Cambridge [England] ; New York: Cambridge University Press.

Hwang, J., Brennan, S. E., & Huffman, M. K. (2015). Phonetic adaptation in non-native spoken dialogue: Effects of priming and audience design. *Journal of Memory and Language*, *81*, 72–90.

Kim, M. (2012). *Phonetic accommodation after auditory exposure to native and nonnative speech*. Northwestern University.

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, *2*(1). https://doi.org/10.1515/labphon.2011.004

Kraljic, T., Brennan, S. E., & Samuel, A. G. (2008). Accommodating Variation: Dialects, Idiolects, and Speech Processing. *Cognition*, *107*(1), 54–81. https://doi.org/10.1016/j.cognition.2007.07.013

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: Tests in Linear Mixed Effects Models*. Retrieved from https://CRAN.R-project.org/package=lmerTest

Lametti, D. R., Krol, S. A., Shiller, D. M., & Ostry, D. J. (2014). Brief Periods of Auditory Perceptual Training Can Determine the Sensory Targets of Speech Motor Learning. *Psychological Science*, *25*(7), 1325–1336.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343.

Lisker, L. (1957). Closure Duration and the Intervocalic Voiced-Voiceless Distinction in English. *Language*, *33*(1), 42. https://doi.org/10.2307/410949

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English/r/and /l/. III. Long-term retention of new phonetic categories. *The Journal of the Acoustical Society of America*, *96*(4), 2076–2087.

Luce, P. A., & Charles-Luce, J. (1985). Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production. *The Journal of the Acoustical Society of America*, *78*(6), 1949–1957.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*(2), 204–238.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, *119*(4), 2382. https://doi.org/10.1121/1.2178720

Pickering, M. J., & Branigan, H. P. (1999). Syntactic priming in language production. *Trends in Cognitive Sciences*, *3*(4), 136–141.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(02), 169–190.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(04), 329–347. https://doi.org/10.1017/S0140525X12001495

R Core Team. (2013). *R: A language and environment for statistical computing.* Vienna, Austria.: R Foundation for Statistical Computing,. Retrieved from http://www.R-project.org/.

Raphael, L. J. (1972). Preceding Vowel Duration as a Cue to the Perception of the Voicing Characteristic of Word-Final Consonants in American English. *The Journal of the Acoustical Society of America*, *51*(4B), 1296–1303. https://doi.org/10.1121/1.1912974

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America*, *88*(1), 97.

Trofimovich, P., & Kennedy, S. (2014). Interactive alignment between bilingual interlocutors: Evidence from two information-exchange tasks. *Bilingualism: Language and Cognition*, *17*(04), 822–836.

Weber, A., Betta, A. M. D., & McQueen, J. M. (2014). Treack or trit: Adaptation to genuine and arbitrary foreign accents by monolingual and bilingual listeners. *Journal of Phonetics*, *46*, 34–51. https://doi.org/10.1016/j.wocn.2014.05.002

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1–25.

CHAPTER 6
GENERAL DISCUSSION

This chapter summarizes the main empirical findings of the preceding chapters and discusses the findings in a broader perspective. It also reviews potential limitations and directions for future research.

EMPIRICAL FINDINGS

This thesis investigated prediction in L1 and L2 auditory comprehension, as well as phonetic adaptation in speech production after comprehension. Many studies have found evidence for predictive processing in the L2 (Chambers & Cooke, 2009; Foucart, Ruiz-Tada, & Costa, 2015; Hopp, 2013, 2015; Ito, Corley, & Pickering, 2017; Ito, Pickering, & Corley, 2018), but others have found weaker prediction effects in the L2 than in the L1, or no prediction at all (Hopp, 2015; Ito et al., 2018; Kaan, Kirkham, & Wijnen, 2014; Martin et al., 2013; Sagarra & Casillas, 2018; van Bergen & Flecken, 2017). Prediction in L2 thus seems to be limited to particular situations. L2 disadvantages have been found mostly for prediction of word form and prediction based on syntactic information. In this thesis, we have focused on semantic prediction. We first tested whether prediction based on verb semantics occurred in L2 and whether there was a disadvantage compared to L1 (in bilinguals and in monolinguals) (**CHAPTER 2)**. We then focused on more naturalistic (longer and more variable) sentences and tested whether there was an L2 disadvantage in prediction of upcoming referents competitors based on lexical-semantic information from the sentence context in **CHAPTER 3**. In this chapter, we also studied pre-activation of semantic competitors with variable strengths of relatedness to the targets for prediction. We then focussed on what mechanisms underlie the L2 semantic disadvantage, when it occurs, in **CHAPTER 4**. Here, we manipulated cognitive load and stimulus presentation rate to test whether limited availability of cognitive resources and slower processing speed in L2 can account for L2 prediction disadvantages. Finally, in **CHAPTER 5** we investigated whether listening to a native speaker of English leads to adaptations in English as L2 speech production.

Lexical-semantic prediction in L2

Using two versions of the visual world paradigm, we investigated whether semantic prediction during comprehension was weaker in the L2 than in the L1. In **CHAPTER 2** looking behaviour during exposure to sentences with constraining and neutral verbs was compared (e.g. *Mary reads a letter* versus *Mary steals a letter* with a display containing objects that were all stealable but of which only one was readable). Dutch-English bilinguals looked at the target object (before hearing the target word could affect fixations) more often in the constraining than in the neutral condition in Dutch and in English, and monolinguals did so in English. This prediction effect did not differ between L1 and L2 within bilinguals, nor between monolinguals in L1 and bilinguals in L2. In bilinguals, the effect became significant in a few time bins later in both L1 and L2 compared to monolinguals. The lack of an effect in earlier bins in bilinguals may have been due to insufficient power in the dataset for the bilinguals to detect an effect (as monolinguals listened to all stimulus sentences in English whereas the bilinguals listened to half the sentences in L1 and half the sentences in L2). If there is indeed a slight bilingual delay, the finding is consistent with the weaker links hypothesis, which suggests that a bilingual necessarily divides language use between two languages and therefore links between semantics and phonology are weaker in bilinguals than in monolinguals. Weaker links could delay lexical access and thereby prediction effects in bilinguals compared to monolinguals. However, overall, the study shows that there was no significant L2 disadvantage in semantic prediction of upcoming referents based on verb semantics in simple SVO sentences.

In **CHAPTER 3 and 4**, predictive processing was assessed by comparing fixations to predicted targets (or semantic competitors) and fixations to unrelated objects in sentences with variable cloze probabilities. The results of the study reported in **CHAPTER 3** again showed that bilinguals predict upcoming semantic information based on the lexical-semantic sentence context in L1 and L2, but now both target and competitor pre-activation was weaker in the L2 than in the L1. Importantly, pre-activation of the competitor was not only modulated by language, but also by the semantic distance between the target for prediction and the semantic competitor. Specifically, the effect of semantic distance on predictive looking behaviour was weaker and started later in L2 than in L1. The sentences were of more variable length and syntactic complexity than the sentences used for the experiment reported in **CHAPTER 2**, leading to more naturalistic stimuli. The display preview time was 500ms before target onset in **CHAPTER 3 and 4**, versus 2200 ms before sentence onset in **CHAPTER 2**. Thus, in more challenging conditions, L2 disadvantages in semantic prediction do arise. Activation of targets for prediction is weaker, and there is slower and/or weaker spreading of semantic activation in L2 than in L1.

The L2 disadvantage found in **CHAPTER 3**, may be caused by weaker links between representations in L2, but also (or as a consequence) by more resource intensive (e.g., Francis & Gutiérrez, 2012) or slower processing in L2 (e.g., Shook, Goldrick, Engstler, & Marian, 2015). Previous research indeed showed that working memory and processing speed modulate predictive eye-movements (Huettig & Janse, 2016). Therefore, in **CHAPTER 4**, we explored these possibilities and tested whether limited availability of cognitive resources or limited processing speed underlie the L2 disadvantage in semantic prediction. The same materials and paradigm as in **CHAPTER 3** were used, but in two experiments we added manipulations of cognitive load (Experiment 1) and processing speed (Experiment 2).

In both experiments we replicated the L2 disadvantage in semantic prediction of targets in the no load/ non-speed-manipulated trials (half the trials). Also, competitor pre-activation was again weaker in L2 than in L1 in both experiments, albeit only in the full data set that had the same amount of trials as the design in **Chapter 3** (including load/speed-manipulated trials). The finding that the modulation of competitor pre-activation by semantic distance with the target was weaker in L2 than in L1 (**CHAPTER 3**) was not replicated in either experiment reported in **CHAPTER 4**, in neither the no-load/non-speed-manipulated trial data sets nor the full data sets. Note that, in the first, power may have been insufficient given the fact that the dataset contained only half the trials of **CHAPTER 3**. In the latter (full) data set, predictive processing was affected by cognitive load and speed and that may have interfered with the semantic distance effect. Either way, the effect needs to be treated with caution, as we failed to replicate it twice. A direct replication needs to be performed to test whether the modulation of competitor prediction by language and semantic distance found in **CHAPTER 3** is a true effect or not. Importantly though, the L2 disadvantage in prediction of targets and competitors was a robust finding, replicated twice.

The results discussed so far are consistent with other studies investigating semantic prediction in bilinguals. For instance, like the results reported in **CHAPTER 2**, Hopp (2015) showed that L2 listeners predicted upcoming semantic information based on lexical-semantic information from the (SVO) sentence context, like L1 listeners did. Ito, Corley and Pickering (2017) contrasted constraining and neutral verbs in a between-participants visual world study, and found that bilinguals listening in L2 showed anticipatory eye-movements to targets in the constraining compared to the neutral condition, like L1 listeners did. In another visual world paradigm, Ito, Pickering and Corley (2018) presented L1 and L2 listeners with more complex sentences including a predictable target (*The tourists expected rain when the sun went behind the cloud$_{[target]}$, but the weather got better later*), and here the authors found an L2 disadvantage in semantic prediction. Specifically, L2 listeners anticipated the target word (cloud) but they did so later than L1 listeners. This finding is consistent with our finding that there is an L2 disadvantage in semantic prediction in more challenging conditions (**CHAPTER 3 and 4**).

Mediating Factors: Cognitive Load and Processing Speed

CHAPTER 4 focussed on the factors that potentially underlie L1-L2 difference in semantic prediction. One potential factor is the cognitive load associated with L2 processing. L2 processing seems to require more cognitive resources than L1 processing (Abutalebi, 2008; Francis & Gutiérrez, 2012; McDonald, 2006), and as a result comprehenders may lack the cognitive resources required for efficient prediction. In Experiment 1 in **CHAPTER 4**, participants were asked to remember 9 non-word syllables on half the trials (in a blocked design). The results showed that an additional cognitive load attenuated predictive looking behaviour to target objects in both the L1 and the L2, but contrary to our expectations, the effect of load was larger in the L1. If L2 comprehension indeed requires more resources than L1 comprehension, predictive processing should be particularly affected by an additional cognitive load in L2. However, the opposite effect was found. Prediction of semantic competitors was not affected by cognitive load in either language, suggesting that prediction of semantic competitors was largely established via routes that were not resource-intensive, such as via low-level lexical associations.

In the visual world paradigm, working memory may serve as a mechanism to hold and bind visual and linguistic representations (stored in long-term memory) to each other and to the current situation (including object location) (Huettig & Janse, 2016; Huettig, Olivers, & Hartsuiker, 2011). Thus the additional cognitive load in **CHAPTER 2** may have hindered this mechanism in which linguistic and visual representations are bound to 'the there and then, or when planning things, to the there and then' (Huettig, Olivers, et al., 2011, p. 143). Also, cognitive resources are required for prediction-by-production, because all stages of production require resources (Pickering & Gambi, 2018). Thus, if target prediction came about mainly via prediction-by-production, an additional cognitive load should indeed attenuate or delay prediction. The particular cognitive load used in this paradigm, may have affected prediction-by-production in an alternative way. Namely, the participants were remembering syllables during half of the trials. To do so, the participants may have been repeating the syllables using inner speech and thus engaging the production system.

L2 processing is slower compared to L1 processing (Cop, Keuleers, Drieghe, & Duyck, 2015; Hahne, 2001; Moreno, Rodríguez-Fornells, & Laine, 2008). As prediction-by-production is likely to require time, slower processing may be another factor underlying the difference between prediction in the L1 and the L2. In Experiment 2 (**CHAPTER 4**) half of the auditory stimuli were presented at a higher rate in L1 and at a lower rate in L2. The results showed that fast speech attenuated prediction of target objects (but not competitor objects) in L1, and slower speech enhanced prediction of competitor objects (but not target objects) in L2. This experiment thus confirms that resources and time can be implicated in predictive processing. Processing speed particularly affected prediction of targets in L1. Presentation rate did not affect prediction of competitors in L1, but prediction of competitors was enhanced by slower presentation rate. Thus, whereas spreading semantic activation to competitors seems to require no (or few) cognitive resources in either language, available time did seem to play a role in L2. The findings are consistent with previous research showing that cognitive load and processing speed affect anticipatory eye-movements (Huettig & Janse, 2016; Ito et al., 2017).

Adaptation of L2 Speech Production after Speech Perception

In **CHAPTER 2-4**, we found an L2 disadvantage in semantic prediction in more demanding situations. Recent accounts of prediction comprehension assume that speech production plays an important role in predictive language processing (Dell & Chang, 2013; Gambi & Pickering, 2013; Pickering & Gambi, 2018; Pickering & Garrod, 2013). In one account, Gambi and Pickering (2013) propose that listeners use production mechanisms to covertly imitate incoming speech and to create forward model predictions of the input. As the production system is involved in this process, adaptations in perception based on incorrect predictions can also affect subsequent speech production. The extent to which comprehenders rely on forward models (and thus the extent to which the production system is involved during comprehension) may depend on social variables such as perceived similarity with the speaker. This account entails that, prediction using the production system and therefore subsequent adaptation in production may fail if interlocutor pairs are highly dissimilar.

In **CHAPTER 5,** we tested whether listening to speech in an L2 (produced by an L1 speaker of that language) could also affect subsequent speech production. Bilingual participants were asked to read aloud sentences containing target phonemes in L2 before and after exposure to speech produced by a native speaker of the participants' L2. The target phonemes were English /æ/ and word-final /b/. These target sounds were chosen because they are notoriously difficult for native speakers of Dutch. The results showed that bilinguals adapted their pronunciation (preceding vowel length and closure duration) of word-final /b/ after exposure to a native speaker producing sentences containing those phonemes. However, no adaptation of /æ/ was found on any of the measures. The extent to which participants adapted their speech did not depend on the physical presence of the confederate, nor on the amount of trials (0 or 4) between the confederate's last production of the phoneme and the participants'. If adaptation were purely automatic, the participants would not only adapt to the confederate's speech, but also vice versa, the competitor would adapt her pronunciation to the participant(s). However, such an effect was not found consistently here.

The results are in line with the findings of Hwang et al. (2015), who also found adaptation of L2 speakers in L2-L1 dialogue for /b/ (on preceding vowel duration but not closure voicing duration). Hwang et al. studied Korean-English speakers. The finding that adaptation persists over four intervening trials extends the observations of Hwang et al. (2015), who found alignment in L2 speech after immediate priming by the L1 confederate without including a lag between target words. The lack of an effect of social context (confederate presence) was surprising, as a mere picture of a speaker on the screen enhanced adaptation in a (unilingual) study by Babel (2012). Gambi and Pickering (2013) suggest that phonetic adaptation through simulation depends on the allocation of limited attentional resources. Perhaps in our study, due to disadvantages associated with L2 processing, the non-native speakers had less resources available to further adjust their pronunciation of /b/ to that of the confederate's when the confederate was present.

Taken together, the result suggest that bilinguals can indeed adjust L2 production after perception, but adjustments are not made for any phoneme. Whether or not L2 speakers adapt production may depend on the extent to which they perceive a difference between the L2 speaker's and their own speech.

THEORETICAL IMPLICATIONS

L2 Disadvantage

Kaan (2014) suggests that the mechanisms underlying prediction in L1 and L2 are the same. In this view, an L2 disadvantage in prediction during language comprehension may arise depending on factors associated with L2 processing, which also underlie individual differences in native processing. Thus, L1-L2 differences in factors such frequency biases, competing information, quality of representation and resulting increases in required cognitive resources and time may give rise to differences between prediction in L1 and L2.

The finding that semantic prediction occurs in L2 (**CHAPTER 2-4)**, sometimes even to the same extent as in L1 (**CHAPTER 2**) is compatible with the view that the mechanisms underlying prediction in L1 and L2 are not qualitatively different. However, an L2 disadvantage in semantic prediction and spreading activation was found in a more demanding context (**CHAPTER 3-4**). One factor that may have given rise to an L2 disadvantage in prediction is frequency of use. Using the L2 less than the L1 may result in weaker links between phonology and semantics (Gollan, Montoya, Cera, & Sandoval, 2008; Gollan et al., 2011), which may in turn result in weaker or slower retrieval of representations (and their associates) required for efficient prediction. Evidence for this hypothesis was found in **CHAPTER 3**, which showed that in more challenging conditions, L2 disadvantages in semantic pre-activation do arise. Activation of semantic competitors of predictable words depended on the semantic distance between targets and competitors and this effect was slower and/or weaker in L2 than in L1. Consistent with the weaker links hypothesis, this result indicates that spreading of semantic activation is weaker and or slower in L2 than in L1. The finding can also be interpreted in terms of the richness of L2 semantic representations. L2 words may have fewer 'senses' (Finkbeiner, Forster, Nicol, & Nakamura, 2004) or semantic features (Schoonbaert, Duyck, Brysbaert, & Hartsuiker, 2009), which may result in slower or narrower spreading activation between concepts. Note that we failed to replicate the weaker effect of semantic distance in L2 in **CHAPTER 4**, therefore the result and possible interpretations needs to be treated with caution.

Also consistent with the view that the mechanisms underlying prediction in L1 and L2 are essentially the same, processing speed and cognitive load affected semantic prediction in both L1 and in L2. The effects of cognitive load and processing speed were more pronounced in the L1 than in the L2, whereas we expected the effects to be particularly large in L2 because of the increased use of cognitive resources and the delay associated with L2 processing. The finding suggests that these factors do not underlie the L2 disadvantage in semantic prediction found in some studies. One explanation for this unexpected finding could be that language does not (only) affect predictive processing across the board, but that it also has an effect on the weights assigned to different routes to prediction. Specifically, the effect of a cognitive load and processing speed may be larger in L1 because bilinguals rely less on resource and time intensive routes to prediction in L2 than in L1.

Do Bilinguals Rely Less on Resource and Time Intensive Routes to Prediction in L2?

Pickering and Gambi (2018) suggest that resource and time intensive prediction-by-production is the most effective route to prediction, but also that it is optional. Prediction-by-association on the other hand, is less targeted but also relatively automatic (thus not optional) and therefore less affected by available time and resources. The authors proposed that language learners (due to limited availability of resources and slower processing speed) may rely less on resource and time-intensive prediction-by-production. In turn, prediction-by-association should be unaffected by limited resources and processing speed. Our results largely support this hypothesis.

First, evidence for the first part of the hypothesis (less reliance on prediction-by-production in L2) was found in **CHAPTER 3 and 4**. There, we found that prediction of target words was weaker in L2 than in L1. Predictions of target word semantics were based on the sentence content of sentences of variable length and complexity, and therefore we expect predictions to be based on higher-order meaning to a large extent. Pickering and Gambi (2018) suggest that prediction of syntactic information and word-form are more affected in populations with less resources and lower processing speed, whereas prediction of semantics is relatively intact (as prediction-by-production occurs in the same order as actual production). Here we show that prediction of semantics can also be affected in language learners.

The results of the manipulations of cognitive resources and processing speed manipulations largely support the hypothesis that comprehenders engage less in prediction-by-production in L2: Speed and available resources affected prediction of target word semantics in both languages, but the affect was more pronounced in the L1. In other words, bilinguals rely more on resource and time-intensive routes to prediction in L1, and therefore the effect of additional cognitive load and presentation rate is larger in L1.

The evidence for the second part of Pickering and Gambi's hypothesis (2018), which implies that prediction-by-association should not be affected by language, is less straight-forward. The first piece of evidence is provided in **CHAPTER 2** in this dissertation. Here we found that semantic prediction based on the verb in simple SVO sentences (e.g. *Mary reads a letter* vs *Mary steals a letter*) did not differ significantly between the L1 and L2 in bilinguals, and between bilinguals in either language and monolinguals. As predictions in this study were based on the semantics of only one word (in combination with the visual object display), they may well have come about by low level lexical associations (as in prediction-by-association).

Further evidence is provided by the pattern of pre-activation of semantic competitors in **CHAPTER 4**. The pre-activation of semantic competitors is likely to be untargeted, as the average plausibility of the competitors as sentence endings was relatively low. Thus, the activation was likely the result of automatic spreading of semantic activation (as in prediction-by-association). In line with the hypothesis that prediction-by-association is unaffected by resources, we found that pre-activation of semantic competitors was not affected by cognitive resources in either language (Experiment 1 in **CHAPTER 4**). Thus, an additional cognitive load in L2 processing (already associated with higher cognitive load than L1 processing) did not seem to affect competitor pre-activation.

At first glance, the L2 disadvantage in competitor pre-activation found in **CHAPTER 3 and 4** seems to be evidence against the hypothesis that prediction-by-association is intact in L2. However, the results are still in line with Pickering and Gambi's framework (2018), if we assume that one of the two following explanations is correct. First, semantic activation may have spread directly from content words in the sentence to semantic competitors, but activation of competitors may also be a result of target word activation. Thus, the activation of competitors is likely to have come about mainly by the prediction-by-association route to prediction, but higher order information may also have played a role if activation of competitors was indirect (via targets). Therefore, even though activation of competitors was a result of prediction-by-association, competitors were activated less strongly in L2 than in L1 because target word activation was also weaker in L2 than in L1.

Second, language may not affect the extent to which bilinguals engage in prediction-by-association, but language experience likely alters semantic associations between word forms and concepts. For instance, some word combinations may have been encountered less frequently in L2 than in L1. Thus, competitor pre-activation was likely weaker in L2 because of weaker associations in L2 and not because of the extent to which bilinguals engaged in prediction-by-association in L2. This explanation is also in line with the finding that the effect of semantic distance between targets and competitors on prediction was weaker and/or later in L2 than in L1 in **CHAPTER 3**, although this interaction was not replicated in **CHAPTER 4**. Finally, the finding that slower presentation rate enhanced prediction of competitors in L2 but not in L1 is also in line with the hypothesis that semantic associations are weaker in L2, if we assume that spreading semantic association therefore requires more time. This assumption is in contrast with Pickering and Gambi's hypothesis that prediction-by-association does not require time however.

Taken together, the results mostly fit with Pickering and Gambi's theory that language learners may use the resource and time intensive route to prediction less. However, the results can also be interpreted in terms of other multi-mechanism accounts of predictive processing (e.g., Huettig, 2015; Kuperberg, 2007; Pickering & Garrod, 2013), as long as they are compatible with the view that some routes to prediction are more resource intensive than others and that more resource intensive routes are optional.

Implications for Models of Bilingual Speech Comprehension

To our knowledge, the Bilingual Language Interaction Network for Comprehension of Speech (BLINCS) is the only implemented models of bilingual comprehension of speech (Shook & Marian, 2013). It was developed as the auditory equivalent of the dominant BIA(+) models) developed for bilingual reading (Dijkstra & van Heuven, 2002). BLINCS is a connectionist model consisting of an interconnected network of dynamic, self-organising maps. The model assumes four levels of representation: phonological, phono-lexical, ortho-lexical and semantic (Figure 1). A spoken word serves as input for the model and activates the best matching node and neighbouring nodes at the phonological level. The phonological level is shared between languages. Visual information from articulatory lip and mouth movements can influence activation at this level directly. As phonological representations get activated, lexical presentations at the phono-lexical level that match the input receive additional activation, while activation of initial candidates that no longer match the phonological input gradually decay. L1 and L2 representations at the phono-lexical level are integrated but separated in space due to within-language phono-tactic probabilities. When phono-lexical items are activated, nearby items are also activated to some extent, but activated items also inhibit items in close proximity. Activation at the phono-lexical level can also feed back to the phonological level, thereby accounting for effects of lexical knowledge on phoneme activation. Phono-lexical items transfer their activation to the ortho-lexical and semantic levels, where activation also spreads to nearby items as a function of proximity. Importantly, activation at the latter two levels also feeds back to the phono-lexical level. Thus, associated representations at the semantic level can become activated at the phono-lexical level, and this activation can even feed back to the phonological level. Thus, activation at the phono-lexical level depends on activation at the phonological, ortho-lexical and semantic levels. In addition, activation at the semantic level is affected by information from a visual scene, such as in the visual world paradigm. Items in the visual context can boost activation of corresponding semantic nodes. Variability in the bilingual system, such as changes in proficiency or recent exposure, is accounted for by the self-organizing feature of the model, which allows the system to make dynamic changes to the maps at all levels.

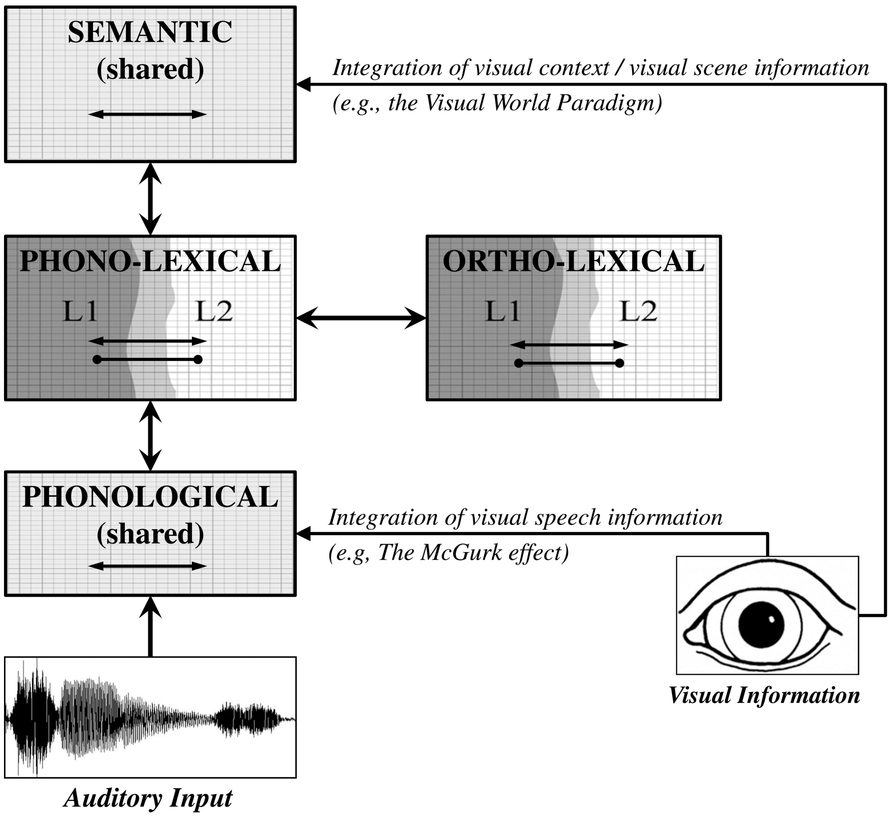


Figure 1. Bilingual Language Interaction Network for Comprehension of Speech (BLINCS) (Shook & Marian, 2013).

Note that the model's mechanisms described above only describe activation flow at the word level. Shook and Marian note that the model is yet to be extended to incorporate prior activation from the linguistic context and effects of expectations based on context information. Therefore, to date, the model does not (yet) imply predictions for the paradigms used **in CHAPTER 2-4,** nor for the observed differences between L1 and L2**.** In order to account for effects of pre-activation due to lower level associative connections between concepts, the model might not need too much adjustment. Basically, resting activation from one or multiple prior incoming words should be allowed to linger while new input comes in, instead of the artificial simplification to reset activation levels for each new incoming word. For instance, upon hearing *Mary knits...* (**CHAPTER 2**), semantic activation from the input *knits* should activate the associated (and therefore nearby) semantic representation of *scarf*, particularly because there is also a scarf present in the concurrent visual display. The semantic activation of *scarf* may feed back to lower levels such as the phono-lexical stage and potentially even the phonological level. Assuming that semantic nodes are mostly language independent, activation should spread to both L1 and L2 lower level representations. This way, during (part of the) verb and the determiner in the spoken sentence *Mary knits a scarf*, the semantic, and perhaps the phono-lexical and phonological nodes for scarf are activated, before the word *scarf* is encountered in the input. Assuming that the semantic representations are language independent, both *scarf* and the Dutch translation equivalent *sjaal* would become activated. Perhaps due to resting activation of *Mary knits* at the phono-lexical level, the English word *scarf*, which would be mapped in closer proximity, would receive more activation than *sjaal*.

In BLINCS, lexical frequency determines the baseline activation of representations at the phono-lexical level, and lateral and between level connections between nodes are strengthened by the frequency with which the relevant representations are active simultaneously. In unbalanced bilinguals, the L2 is usually used less frequently than the L1, and therefore the combinations of word form and semantic representations for a lexical item, but also combinations of two associated lexical items are encountered less frequently in L2. Although this is not discussed explicitly by Shook and Marian (2013), this should result in weaker links between representations within and between levels, and also weaker baseline activation levels in L2. If BLINCS were to be adjusted to incorporate prediction by association, weaker prediction effects are thus expected in L2 than in L1. This is consistent with the finding that pre-activation of semantic competitors was weaker in L2 than in L1 in **CHAPTER 3 and 4.** It is however inconsistent with **CHAPTER 2**, where we found no effect of language on semantic prediction. Perhaps, there was no significant difference between languages here, because the simple associations between verbs and nouns were quite strong and because the visual input provided even more prior activation of the context. Note that in **CHAPTER 3 and 4**, the semantic competitor was not the target for prediction in the sentence, and that the visual scene was presented only 500ms before target word onset (versus 2200 ms before sentence onset).

Incorporating prediction based on higher lever information (such as prediction-by-production) in BLINCS, sensitive to multiple linguistic constraints (e.g. syntactic constraints), is much more challenging. The model needs to be extended to incorporate higher order meaning, and word (pre-) activation should depend on other linguistic features such as syntactic position. Thus, prior sentence context should be allowed to directly activate semantic representations like the visual context can. But, it may also need to directly affect phono-lexical representations (for instance, words in a particular syntactic category should receive more activation). This way, the model would be able to account for the effects of target prediction in **CHAPTER 3 and 4.** The L2 disadvantage in target word prediction in **CHAPTER 3 and 4** can partly be accounted for by weaker baseline activation of L2 words and weaker connections between L2 phono-lexical and semantic representation, like in prediction-by-association. However, the findings in **CHAPTER 4** suggest that bilinguals may engage less in prediction-by-production in L2 than in L1. To account for the finding, the model should also incorporate a way to shift the weights assigned to activation from higher order information, visual information and associative connections, depending on factors such as language use, available cognitive resources and input speed. This way the model would be able to account for weaker reliance on prediction-by-production in L2, and for increased use of non-linguistic contextual cues (e.g. visual input) (Bradlow & Alexander, 2007; Navarra & Soto-Faraco, 2007) .

In **CHAPTER 5** we found that spoken language input cannot only lead to predictions in L2 language comprehension such as in **CHAPTER 2-4**, but also that it can lead to changes in subsequent speech production. So far, we have mainly discussed this finding in terms of Gambi and Pickering's (2013) account of imitation in speech. According to these authors, phonetic imitation is achieved via correction of prediction errors: A listener simulates incoming speech using the production system and generates forward model predictions. Due to the involvement of the prediction system, adaptations based on prediction errors in comprehension can also affect subsequent speech production, depending on contextual factors such as speaker-listener perceived similarity. The finding that adaptation only occurred for word final /b/ and not for /æ/ could be explained in terms of noticing prediction error. Dutch-English bilinguals have difficulty distinguishing /æ/ from their native vowel /ɛ/. If L2 listeners do not notice that a speaker produces a particular phoneme differently than they predicted, they cannot use prediction error to make adaptations in perception and production. BLINCS is a model of comprehension and does not make assumptions about speech production. If the model were to assume parity between production and comprehension, then it could possibly account for adaptations in production based on prior input as well. Specifically, in the case of our Dutch-English unbalanced bilinguals, repeatedly hearing a word such as *mob* with a voiced word-final /b/ spoken by an English native speaker, would strengthen the associations between the representations of the phonemes in *mob*, and the English phono-lexical representation. Thus, in subsequent production, activation of the phono-lexical representation of *mob*, would be more likely to activate the native English phoneme representations at the phonological level (/b/ and not /p/) due to strengthening of the link between the phonological and phono-lexical representations.

Dutch-English bilinguals have a representation of /b/ in L1, although in Dutch it never appears in word-final position. Therefore, the adaptation effect can be explained by strengthening the connections between the phono-lexical representation *mob* and the English native phoneme sequence (because connections between items that are activated together are strengthened). However, the vowel /æ/ does not exist in Dutch. Therefore, initially, when the bilinguals do not have a node for /æ/ yet, or when the links to and from the node are still very weak, the input /æ/ may be mapped to the Dutch vowel /ɛ/ in Dutch-English bilinguals, which would then be the 'best-match unit' for /æ/. When input is mapped to a particular node, "the value of the node is altered to become more similar to the input" (Shook & Marian, 2013). Perhaps then, only with significant input, a separate node would arise for the English phoneme /æ/. Only then, the links between the phoneme /æ/ and phono-lexical representations such as *map* could be promoted by the input. Again, assuming parity between production and comprehension, strengthening of these connections would subsequently lead to adjustments in speech production. Thus, the finding that phonetic adaptation occurred for word-final /b/ but not for /æ/ in Dutch-English bilinguals could be explained using BLINCS as framework, but only if we assume the input /æ/ is initially mapped to the phonological representation /ɛ/, and if there is parity between production and comprehension. This interpretation entails that more input should eventually lead to adaptation for the phoneme /æ/ as well.

LIMITATIONS AND FUTURE RESEARCH

The present dissertation found evidence for an L2 disadvantage in semantic prediction by bilinguals listening to spoken language. The results are also in line with the hypothesis that bilinguals listening to spoken language rely less on resource intensive routes to prediction in L2 than in L1. The studies in **CHAPTER 2-4** were designed to test whether there was a direct effect of language on semantic prediction and whether weaker links, limited resources or slower processing speed played a role in the effect, and not to test whether the relative reliance on prediction-by-production and/or prediction-by-association depended on language. In our design, we cannot be fully certain that prediction of competitor objects came about by prediction-by-association and prediction of target objects by prediction-by-production. Future research could therefore be aimed at investigating whether bilinguals indeed engage in the latter less in L2 than in L1. One study showing consistent results is a visual world study by Peters et al., (2015). In this study, high and low proficient bilinguals listened to SVO sentences (e.g., *The pirate chases the ship*) while they looked at displays featuring agent-related, action-related and unrelated pictures. The low-proficient bilinguals were more likely than high-proficient bilinguals to anticipate locally-coherent action-related distractors (e.g., a cat), suggesting that the low proficient bilinguals' predictions relied more on low-level word associations than high proficient bilinguals' predictions. Peters et al. interpret the results in terms of adaptation to uncertainty, with low proficient bilinguals adapting to inherently higher levels of uncertainty by activating locally coherent distractors more strongly. Another way to investigate directly whether the production system is engaged more in prediction in L1 than in L2 is by testing whether engaging the production system in a secondary task (e.g. repeating a syllable using inner speech; see Martin, Branzi, & Bar, 2018) during comprehension has a larger effect on prediction in L1 than in L2.

In this dissertation, we established that there is an L2 disadvantage in semantic prediction during speech comprehension, at least in some situations. We also found evidence for the hypothesis that prediction in L2 is weaker due to weaker associative connections between concepts in **CHAPTER 3** (although this finding was not replicated in **CHAPTER 4**). However, the L2 disadvantage in semantic prediction may have other loci as well. For instance, in order to generate correct predictions (according to Pickering & Gambi, 2018), one needs to derive speaker intention (the probability that *I want to give you a ...* ends in *car* for instance, is strongly dependent on who says it and in what situation it is said). However, deriving speaker intention may be more difficult in L2 than in L1. This could be due to the comprehender sharing less background knowledge with the speaker, but also due to the L2 comprehender relying more on non-linguistic context information in this process than an L1 comprehender (Bradlow & Alexander, 2007; Navarra & Soto-Faraco, 2007). Future research is needed to test whether deriving speaker intention is more difficult in L2 than in L1 and whether L2 comprehenders rely more on non-linguistic information in the process.

One downside of the visual world paradigm employed in **CHAPTER 2-4**, is that there is by definition a visual context present with the auditory stimuli. Thus, spoken language may activate the objects depicted on the screen, but the depicted objects may also lead to activation of linguistic representations, especially at longer preview intervals (Huettig, Rommers, & Meyer, 2011). On the other hand, in real life conversation, there is almost always a visual context present, and objects or events in the visual context are often referred to in the conversation. Therefore, the presence of a visual context does not eliminate the ecological validity of the experimental paradigm.[1] Note however, that in experiments on predictive processing, the presence of a limited number of visual

[1] See Eichert, Peeters, and Hagoort (2018) for evidence of anticipatory eye-movements in a more realistic setting (using virtual reality).

objects combined with an auditory sentence be a highly constraining context, and the presence of the target object for prediction on the screen (particularly over several trials) may strongly encourage prediction. Therefore, visual world experiments on prediction during comprehension often show what listeners are capable of doing, not what they actually do in an average conversation (Huettig, Rommers, et al., 2011; and see Salverda, Brown, & Tanenhaus, 2011 for a discussion of the relevance of participants' goals in the visual world paradigm).

Throughout this dissertation we have discussed that prediction is the result of linguistic and non-linguistic experience. In order to make sure that any differences in auditory processing between languages were the result of language experience only, we always used within-subject designs. This also eliminated possible confounding effects of individual cognitive differences (such as in working memory, executive function, processing speed et cetera). On the other hand, we always tested participants from the same participant pool and age group with a relatively high proficiency level. As bilinguals vary in terms of proficiency, age of acquisition, context of acquisition, L1-L2 relatedness, life experience and many other factors, the extent to which (and the way in which) prediction is affected in L2 may also vary. L2 proficiency did not affect semantic prediction in any of our studies, but the participants' proficiency levels may not have been variable enough to detect an effect. Thus, to be able to generalize to other bilingual populations, a large-scale study with a variable sample in which individual differences are assessed should be performed.

In **CHAPTER 5**, we found that bilinguals adapted speech production after perception of speech in their L2 (pronounced by an L1 speaker) for one target phoneme and not for the other. To test whether adaptation indeed fails to occur when the difference between the predicted and perceived phoneme is not noticed (and/or because they have no representation of the L2 phoneme), future research could include a phoneme categorisation task to test where L2 listeners place the boundary between a non-native phoneme (such as /æ/) from a similar native one (such as /ɛ/). The hypothesis would be that if listeners perceive the difference between their representation of /æ/ and the L1 speaker's production, they adapt their expectations of what /æ/ sounds like in perception. This should result in a shift in the /æ/-/ɛ/ boundary in perception and a subsequent adaptation in production. Perceptual adaptation combined with adaptation in production would be evidence for the assumption that there is parity between production and comprehension.

In **CHAPTER 5** we also noted that adaptations in speech production due to speech perception could serve as a useful learning strategy in L2 (see also, Costa, Pickering, & Sorace, 2008). If so, adaptations should be relatively long-term, given sufficient exposure. In **CHAPTER 5** we found adaptation effects immediately after perception and with 4 intervening trials (after prior exposure to 30 instances of the target phoneme). However, a study testing whether these effects persist at longer time intervals would be of added value.

Finally, this dissertation focussed mainly on semantic prediction in the auditory modality. As we have noted earlier, predicting upcoming information based on linguistic context may well be more difficult in the auditory modality, because spoken language is inherently more variable than written language, and because of speech unfolding in time rather than in space (like writing). In addition, cross-linguistic interference may be larger in the auditory modality in bilinguals due to difficulty distinguishing native from non-native phonemes. On the other hand, spoken language contains many cues that written language does not. For instance, prosody may signal intended irony or speaker mood, and hesitation affects word integration (Corley, MacGregor, & Donaldson, 2007). These additional cues may also facilitate predictive processing in L2 speakers. An interesting line of research would be to directly compare prediction in the auditory and visual modality in L1 and L2, or to compare prediction in auditory sentences with and without particular prosodic cues for interpretation.

CONCLUSION

The four empirical chapters in this dissertation contribute to the literature on prediction during language comprehension, and to the literature on bilingual spoken language comprehension and production. First, we show that bilinguals can predict semantic information based on lexical-semantic information from the sentence context in the L2, even to the same extent as in the L1. In more demanding contexts, with longer and more variable sentences an L2 disadvantage was found. This disadvantage was replicated in two additional experiment using the same stimuli and design. Second, we provide evidence that slower spreading semantic activation may play a role in the L2 disadvantage in semantic prediction. Weaker links between word form and semantics due to less exposure to L2 might explain this finding. Third, cognitive load and processing speed do not seem to underlie the L2 disadvantage in semantic prediction directly, but these (and potentially other factors) may cause a strategy shift, with bilinguals relying more on a resource intensive route to prediction in L1 than in L2. Cognitive load and processing speed did affect prediction in both L1 and L2, consistent with the view that the mechanisms underlying prediction in L1 and L2 are the same. Finally, listening to speech in the L2 (spoken by an L1 speaker) can result in subsequent phonetic adaptations in production, but not for any phoneme.

REFERENCES

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica*, *128*(3), 466–478. https://doi.org/10.1016/j.actpsy.2008.03.014

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, *40*(1), 177–189. https://doi.org/10.1016/j.wocn.2011.09.001

Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America*, *121*(4), 2339. https://doi.org/10.1121/1.2642103

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 1029–1040. https://doi.org/10.1037/a0015901

Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015). Frequency effects in monolingual and bilingual natural reading. *Psychonomic Bulletin & Review*, *22*(5), 1216–1234. https://doi.org/oi: 10.3758/s13423-015-0819-2

Corley, M., MacGregor, L. J., & Donaldson, D. I. (2007). It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition*, *105*(3), 658–668. https://doi.org/10.1016/j.cognition.2006.10.010

Costa, A., Pickering, M. J., & Sorace, A. (2008). Alignment in second language dialogue. *Language and Cognitive Processes*, *23*(4), 528–556. https://doi.org/10.1080/01690960801920545

Dell, G. S., & Chang, F. (2013). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B:*

Biological Sciences, *369*(1634), 20120394–20120394. https://doi.org/10.1098/rstb.2012.0394

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(03). https://doi.org/10.1017/S1366728902003012

Eichert, N., Peeters, D., & Hagoort, P. (2018). Language-driven anticipatory eye movements in virtual reality. *Behavior Research Methods*, *50*(3), 1102–1115. https://doi.org/10.3758/s13428-017-0929-z

Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language*, *51*(1), 1–22. https://doi.org/10.1016/j.jml.2004.01.004

Foucart, A., Ruiz-Tada, E., & Costa, A. (2015). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition*, 1–7. https://doi.org/10.1017/S1366728915000486

Francis, W. S., & Gutiérrez, M. (2012). Bilingual recognition memory: Stronger performance but weaker levels-of-processing effects in the less fluent language. *Memory & Cognition*, *40*(3), 496–503. https://doi.org/10.3758/s13421-011-0163-3

Gambi, C., & Pickering, M. J. (2013). Prediction and imitation in speech. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00340

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209. https://doi.org/10.1037/a0022256

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, *30*(3), 251–266. https://doi.org/10.1023/A:1010490917575

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*(1), 33–56. https://doi.org/10.1177/0267658312461803

Hopp, H. (2015). Semantics and Morphosyntax in Predictive L2 Sentence Processing. *International Review of Applied Linguistics.* https://doi.org/10.1515/iral-2015-0014

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459

Huettig, F., Olivers, C. N. L., & Hartsuiker, R. J. (2011). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Acta Psychologica*, *137*(2), 138–150. https://doi.org/10.1016/j.actpsy.2010.07.013

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*(2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003

Hwang, J., Brennan, S. E., & Huffman, M. K. (2015). Phonetic adaptation in non-native spoken dialogue: Effects of priming and audience design. *Journal of Memory and Language*, *81*, 72–90. https://doi.org/10.1016/j.jml.2015.01.001

Ito, A., Corley, M., & Pickering, M. J. (2017). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, *21*(2), 251–264. https://doi.org/10.1017/S1366728917000050

Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, *98*, 1–11. https://doi.org/10.1016/j.jml.2017.09.002

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? In *Linguistic Approaches to Bilingualism* (pp. 257–282).

Kaan, E., Kirkham, J., & Wijnen, F. (2014). Prediction and integration in native and second-language processing of elliptical structures. *Bilingualism: Language and Cognition*, 1–18. https://doi.org/10.1017/S1366728914000844

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Research*, *1146*, 23–49. https://doi.org/10.1016/j.brainres.2006.12.063

Martin, C. D., Branzi, F. M., & Bar, M. (2018). Prediction is Production: The missing link between language production and comprehension. *Scientific Reports*, *8*(1), 1079. https://doi.org/10.1038/s41598-018-19499-4

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*(3), 381–401. https://doi.org/10.1016/j.jml.2006.06.006

Moreno, E. M., Rodríguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics*, *21*(6), 477–508. https://doi.org/10.1016/j.jneuroling.2008.01.003

Navarra, J., & Soto-Faraco, S. (2007). Hearing lips in a second language: visual articulatory information enables the perception of second language sounds. *Psychological Research*, *71*(1), 4–12. https://doi.org/10.1007/s00426-005-0031-5

Peters, R., Grüter, T., & Borovsky, A. (2015). Anticipatory and Locally Coherent Lexical Activation Varies as a Function of Language Proficiency. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 1865–1870). Austin, TX.

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: a theory and review. *Psychological Bulletin*. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(04), 329–347. https://doi.org/10.1017/S0140525X12001495

Sagarra, N., & Casillas, J. V. (2018). Suprasegmental information cues morphological anticipation during L1/L2 lexical access. *Journal of Second Language Studies*, *1*(1), 31–59. https://doi.org/10.1075/jsls.17026.sag

Salverda, A. P., Brown, M., & Tanenhaus, M. K. (2011). A goal-based perspective on eye movements in visual world studies. *Acta Psychologica*, *137*(2), 172–180. https://doi.org/10.1016/j.actpsy.2010.09.010

Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2009). Semantic and translation priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition*, *37*(5), 569–586. https://doi.org/10.3758/MC.37.5.569

Shook, A., Goldrick, M., Engstler, C., & Marian, V. (2015). Bilinguals Show Weaker Lexical Access During Spoken Sentence Comprehension. *Journal of Psycholinguistic Research*, *44*(6), 789–802. https://doi.org/10.1007/s10936-014-9322-6

Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for Comprehension of Speech. *Bilingualism: Language and Cognition*, *16*(02), 304–324. https://doi.org/10.1017/S1366728912000466

van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language*, *92*, 26–42. https://doi.org/10.1016/j.jml.2016.05.003

CHAPTER 7
ENGLISH SUMMARY

Bilingualism is not a rare phenomenon. It has been estimated that more than half the people in Europe have knowledge of two or more languages (European Commission, 2012). A second language (L2) is also increasingly used in higher education. Thus, in order to advance our understanding of human language processing, we cannot built only on research on the monolingual brain. Theories on the mechanisms involved in language processing need to be extended for the bilingual case. In this dissertation, we focused on spoken language comprehension in the native and non-native language. We listen to speech on a daily basis, and speech comprehension is usually effortless in our native language (L1). In a second language (L2) on the other hand, speech comprehension is often slower, more effortful, and we sometimes make mistakes in the proces (Hahne, 2001; Moreno, Rodríguez-Fornells, & Laine, 2008; Schmidtke, 2016; Weber & Broersma, 2012).

One key mechanism thought to support L1 language processing is prediction of upcoming information. Prediction can speed up processing, help the comprehender to deal with variability and ambiguity in speech and it can help the listener determine when to start a response in dialogue (e.g., Kutas, DeLong, & Smith, 2011; Van Berkum, 2010). This mechanism could be particularly useful in L2 comprehension, as adaptations of expectations due to prior prediction errors might aid language learning (Dell & Chang, 2013). On the other hand, disadvantages associated with L2 processing may hinder efficient prediction in L2. For instance, increased lexical competition in L2 (Weber & Broersma, 2012), weaker representations (Kaan, 2014) and weaker links between word form and semantics in L2 (Gollan, Montoya, Cera, & Sandoval, 2008; Gollan et al., 2011) may slow down lexical access, or lead to an increase in the resources required. Slower or more effortful retrieval of linguistic representations may hinder the construction of higher-level meaning used to predict an upcoming word, or the retrieval of the to be predicted word itself. Some recent studies have found evidence for predictive processing in the L2 (e.g., Chambers & Cooke, 2009; Foucart, Ruiz-Tada, & Costa, 2015; Hopp, 2015; Ito, Corley, & Pickering, 2017), but others have found weaker prediction effects in L2 than in L1 (e.g., Hopp, 2015; Ito, Pickering, & Corley, 2018; Kaan, Kirkham, & Wijnen, 2014; Martin et al., 2013). Prediction in the L2 thus seems to be limited to particular situations. In this thesis we have studied if and when an L2 disadvantage in prediction arises.

In **CHAPTER 2** we investigated whether bilinguals anticipated upcoming referents based on information extracted at the verb in their L1 and L2, in simple sentences, and whether bilinguals predicted to a similar extent in both languages. In addition, a monolingual control group was tested to see whether any differences between prediction in L1 and L2 were due to the language manipulation (English vs. Dutch) or due to language status (L1 vs. L2). A visual world paradigm was used, in which looking behaviour during exposure to L1 and L2 sentences with constraining and neutral verbs was compared (e.g. *Mary reads a letter* versus *Mary steals a letter* with a display containing objects that were all stealable but of which only one was readable). Dutch-English bilinguals looked at the target object more often in the constraining than in the neutral condition in Dutch and in English, before hearing the target word could affect fixations. Monolinguals did so in English. This prediction effect did not differ between L1 and L2 within bilinguals, nor between monolinguals in L1 and bilinguals in L2. Although in this paradigm predictions could theoretically be generated based on higher level information (e.g. combined meaning of *Mary knits*), it is also likely that association-based mechanisms were involved to a large extent. For instance, semantic activation from the verb *knits* may have spread to the associated concept *scarf*. The findings are consistent with Pickering and Gambi's (2018) account of prediction. This account assumes two routes to prediction: prediction-by-production and prediction-by-association. The first route is the most effective but it dependson available resources and time. The second route is not optional and relatively automatic. The authors assume that, given its automatic nature, prediction-by-association should be mostly intact in populations with limited resources such as L2 comprehenders.

Then, in **CHAPTER 3**, we studied semantic prediction in bilinguals in a more fine-grained way, again using the visual world paradigm. We hypothesized that subtle differences between semantic prediction in L1 and L2 should arise in more challenging conditions. In that case prediction is likely to require more resources, possibly unavailable in L2. The stimulus sentences were more variable, longer and syntactically more complex than the ones used in **CHAPTER 2**. Predictions of the sentence-final target word could be based on the combined higher-level meaning of the words in the sentence. The picture preview time was very short, to avoid visual priming of the target for prediction. We also tested whether pre-activation via spreading of semantic activation differed between L1 and L2 by adding a condition in which the diplay included a semantic competitor (e.g., *arm* for target *leg*) instead of the target for prediction. Consistent with **CHAPTER**

2, **CHAPTER 3** showed that bilinguals predicted upcoming semantic information based on the lexical-semantic sentence context in L1 and L2, but both target and competitor pre-activation was weaker in the L2 than in the L1. Importantly, pre-activation of the competitor was not only modulated by language, but also by the semantic distance between the target for prediction and the semantic competitor. Specifically, the effect of semantic distance on predictive looking behaviour was weaker and started later in L2 than in L1. Thus, in more challenging conditions, L2 disadvantages in semantic prediction do arise. Slower or weaker spreading semantic activation in L2 may be a result of weaker links between word form and semantics due to less practice of the L2 than the L1 (e.g. Gollan et al., 2008).

Differences between semantic prediction in L1 and L2 are not found consistently, and it remains unclear what mechanisms underlie the difference, when it is found. **CHAPTER 4** therefore focused on the factors that potentially underlie the L1-L2 difference in semantic prediction. One potential factor is the cognitive load associated with L2 processing (e.g., Francis & Gutiérrez, 2012; McDonald, 2006). Pickering and Gambi (2018) have suggested that prediction-by-production may be impaired in populations with limited availability of resources, as this route to production requires resources and time. We therefore hypothesized that if prediction in L2 is weaker because of the higher cognitive load associated with L2 processing, an additional cognitive load would be particularly detrimental for prediction in L2. The second potential factor, processing speed, was chosen using a similar line of reasoning. L2 processing is slower compared to L1 processing (Moreno et al., 2008; Shook, Goldrick, Engstler, & Marian, 2015). As prediction-by-production requires time, slower processing may underlie the difference between prediction in the L1 and the L2. In two experiments we manipulated cognitive load (by asking participants to remember 9 syllables on half the trials) and processing speed (by speeding up half the spoken stimuli in L1 and slowing them down in L2) to test these hypotheses. This study used the same materials as **CHAPTER 3** and it was therefore also an attempt to replicate the findings of our previous experiment.

First, in both experiments we replicated the L2 disadvantage in semantic prediction of tagets in the no load/ non-speed-manipulated trials (half the trials). Competitor pre-activation was also weaker in L2 than in L1 in both experiments (like in **CHAPTER 3**), but only in the full data set (all trials, including load/speed-manipulated trials). The finding in **CHAPTER 3** that the modulation of competitor pre-activation by semantic distance with the target was weaker in L2 than in L1 was not replicated in either experiment reported in **CHAPTER 4**. In Experiment 1 in

CHAPTER 4, an additional cognitive load attenuated predictive looking behaviour to target objects in both the L1 and the L2, but contrary to our expectations, the effect of load was larger in the L1. In Experiment 2 (**CHAPTER 4**), fast speech attenuated prediction of targets in L1, but slower speech did not enhance prediction targets in L2. These experiments confirm that resources and processing speed can be implicated in predictive processing (Huettig & Janse, 2016; Ito et al., 2017). A cognitive load and processing speed particularly affected prediction of targets in L1. One explanation for this unexpected finding could be that L2 does not affect predictive processing directly, but that it has an effect on the weights assigned to different routes to prediction. Specifically, the effect of a cognitive load and of processing speed may be larger in L1 because bilinguals rely more on resource intensive routes to prediction in L1 than in L2. Prediction of semantic competitors was not affected by cognitive load in either language, suggesting that prediction of semantic competitors was largely established via routes that were not resource-intensive, such as via low-level lexical associations. Presentation rate did not effect prediction of competitors in the L1, but slower rate did enhance prediction of competitors in L2. Thus, whereas spreading semantic activation to competitors seems to require no (or few) cognitive resources in either language, available time did seem to play a role in L2. The finding that processing speed and cognitive load affected semantic prediction in L1 and in L2 is again consistent with Kaan's proposal that the mechanisms underlying prediction in L1 and L2 are essentially the same (2014).

Recent prominent accounts of predictive language procesing assume representational parity between comprehension and production (e.g., Pickering & Gambi, 2018; Pickering & Garrod, 2013), which entails that listening to speech may affect subsequent speech production. In **CHAPTER 5** we tested whether adaptation occurred in L2. Dutch-English bilingual participants read aloud sentences containing target phonemes /æ/ and word-final /b/ in L2, before and after exposure to speech produced by a native speaker English. Bilinguals adapted their pronunciation of word-final /b/ after exposure to an L1-speaker of Egnlish producing sentences containing those phonemes. However, no adaptation of /æ/ was found on any of the measures. The extent to which participants adapted their speech did not depend on the physical presence of the conferate, nor on the amount of trials between the confederate's last production of the phoneme and the participants'. If adaptation were purely automatic, the participants would not only adapt to the confederate's speech, but also vice versa, the competitor would adapt her pronunciation to the participant(s). However, such an effect was not found consistently here. Taken together, the result suggest that bilinguals can indeed adjust L2 production after perception, but adjustments are not made for any phoneme.

The empirical chapters in this dissertation contribute to the literature on prediction during language comprehension, and to the literature on bilingual spoken language comprehension and production. First, we show that bilinguals can predict semantic information based on lexical-semantic information from the sentence context in the L2, even to the same extent as in their L1. In more demanding contexts, an L2 disadvantage was found when comparing prediction in both languages in the same individuals. Second, we provide evidence that spreading semantic (pre-)activation is slower and/or weaker in L2. Weaker links between word forms and semantic due to less exposure to L2 might explain this finding. Third, cognitive load and processing speed do not seem to underlie the L2 disadvantage in semantic prediction directly, but these and potentially other factors may cause a strategy shift, with bilinguals relying more on a resource intensive route to prediction in L1 than in L2. Cognitive load and processing speed did affect prediction in both L1 and L2, consistent with the view that the mechanisms underlying prediction in L1 and L2 are the same. Finally, listening to speech in the L2 (spoken by an L1 speaker) can result in subsequent phonetic adaptations in production, but not for any phoneme.

REFERENCES

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 1029–1040. https://doi.org/10.1037/a0015901

Dell, G. S., & Chang, F. (2013). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120394–20120394. https://doi.org/10.1098/rstb.2012.0394

European Commission. (2012). Europeans and their languages. Retrieved from http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_386_en.pdf

Foucart, A., Ruiz-Tada, E., & Costa, A. (2015). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition*, 1–7. https://doi.org/10.1017/S1366728915000486

Francis, W. S., & Gutiérrez, M. (2012). Bilingual recognition memory: Stronger performance but weaker levels-of-processing effects in the less fluent language. *Memory & Cognition*, *40*(3), 496–503. https://doi.org/10.3758/s13421-011-0163-3

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209. https://doi.org/10.1037/a0022256

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, *30*(3), 251–266. https://doi.org/10.1023/A:1010490917575

Hopp, H. (2015). Semantics and Morphosyntax in Predictive L2 Sentence Processing. *International Review of Applied Linguistics.* https://doi.org/10.1515/iral-2015-0014

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459

Ito, A., Corley, M., & Pickering, M. J. (2017). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, *21*(2), 251–264. https://doi.org/10.1017/S1366728917000050

Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, *98*, 1–11. https://doi.org/10.1016/j.jml.2017.09.002

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, *4*(2), 257–282. https://doi.org/10.1075/lab.4.2.05kaa

Kaan, E., Kirkham, J., & Wijnen, F. (2014). Prediction and integration in native and second-language processing of elliptical structures. *Bilingualism: Language and Cognition*, 1–18. https://doi.org/10.1017/S1366728914000844

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 190–207).

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*(3), 381–401. https://doi.org/10.1016/j.jml.2006.06.006

Moreno, E. M., Rodríguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics*, *21*(6), 477–508. https://doi.org/10.1016/j.jneuroling.2008.01.003

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: a theory and review. *Psychological Bulletin*. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(04), 329–347. https://doi.org/10.1017/S0140525X12001495

Schmidtke, J. (2016). The Bilingual Disadvantage in Speech Understanding in Noise Is Likely a Frequency Effect Related to Reduced Language Exposure. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00678

Shook, A., Goldrick, M., Engstler, C., & Marian, V. (2015). Bilinguals Show Weaker Lexical Access During Spoken Sentence Comprehension. *Journal of Psycholinguistic Research*, *44*(6), 789–802. https://doi.org/10.1007/s10936-014-9322-6

Van Berkum, J. J. A. (2010). The brain is a prediction machine that cares about good and bad—any implications for neuropragmatics. *Italian Journal of Linguistics*, *22*(1), 181–208.

Weber, A., & Broersma, M. (2012). Spoken word recognition in second language acquisition. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Hoboken, NJ, USA: John Wiley & Sons, Inc. Retrieved from http://doi.wiley.com/10.1002/9781405198431

CHAPTER 8
NEDERLANDSE SAMENVATTING

Tweetaligheid is niet zeldzaam. Volgens een schatting van de Europese Unie hebben meer dan de helft van de europeanen kennis van twee of meer talen (European Commission, 2012). Gebruik van een tweede taal is ook in opkomst in het hoger onderwijs. Om onze kennis op het gebied van taalverwerking te vergroten, kunnen we dus niet enkel bouwen op onderzoek over het eentalige brein. Theoriën over taalverwerking moeten uitgebreid worden zodat ze ook toepasbaar zijn op het tweetalige brein. In dit proefschrift richtten we ons met name op het luisteren naar spraak in de eerste en tweede taal. We luisteren dagelijks naar spraak, en spraakverwerking verloopt meestal moeiteloos in onze moedertaal (T1). In onze tweede taal (T2) daarentegen, verloopt spraakverwerking doorgaans trager, kost het meer moeite en maken we regelmatig fouten (Hahne, 2001; Moreno, Rodríguez-Fornells, & Laine, 2008; Schmidtke, 2016; Weber & Broersma, 2012).

Een primair mechanisme dat T1-verwerking lijkt te ondersteunen is het voorspellen van informatie. Door informatie over wat een spreker mogelijk gaat zeggen (zoals een woord of een syntactische categorie) te voorspellen kan de taalverwerking vlotter verlopen. Het kan de luisteraar ook helpen om te gaan met variatie en ambiguïteit van gesproken taal, en bovendien kan het helpen om te plannen wanneer het jouw beurt is om te beginnen spreken (e.g., Kutas, DeLong, & Smith, 2011; Van Berkum, 2010). Het mechanisme zou ook zeer belangrijk kunnen zijn voor T2-verwerkers, aangezien zij hun bekwaamheid in T2 zouden kunnen verbeteren door te leren van voorspellingen die later fout blijken te zijn (Dell & Chang, 2013). Anderzijds zou het voorspellen weleens moeilijker kunnen zijn in de T2, vanwege moeilijkheden inherent aan T2-verwerking. Bijvoorbeeld, toegenomen concurrentie van woordkandidaten (Weber & Broersma, 2012), zwakkere linguïstische representaties (Kaan, 2014) en zwakkere verbindingen tussen woordvormen en semantiek (Gollan, Montoya, Cera, & Sandoval, 2008; Gollan et al., 2011), zouden kunnen leiden tot vertragingen of grotere belasting van het werkgeheugen tijdens T2-verwerking. Het traag of moeizaam activeren van linguïstische representaties zou voor problemen kunnen zorgen bij het construeren van overkoepelende betekenis, zoals van een zin, die nodig is om bijvoorbeeld een woord te kunnen voorspellen. Het zou ook problemen kunnen geven bij het ophalen van de linguïstische representatie van het te voorspellen woord zelf. Een aantal recente onderzoeken heeft evidentie gevonden voor het voorspellen tijdens T2-verwerking (e.g., Chambers & Cooke, 2009;

Foucart, Ruiz-Tada, & Costa, 2015; Hopp, 2015; Ito, Corley, & Pickering, 2017), maar andere onderzoeken vonden juist zwakkere of zelfs geen voorspelling in de T2 (e.g., Hopp, 2015; Ito, Pickering, & Corley, 2018; Kaan, Kirkham, & Wijnen, 2014; Martin et al., 2013). Voorspellen lijkt in T2 dus te zijn voorbehouden aan specifieke situaties. In dit proefschrift onderzochten we of, en zo ja, in welke omstandigheden het voorspellen zwakker is in de T2.

In **HOOFDSTUK 2** onderzochten we of tweetaligen toekomstige informatie voorspelden op basis van de betekenis van het werkwoord in eenvoudige zinnen in de T1 en de T2, en of de mate waarin voorspeld verschilde tussen de talen. Bovendien werd er ook een monolinguale controlegroep onderzocht om na te gaan of een eventueel verschil tussen voorspellen in T1 en T2 verklaard kon worden door een verschil tussen de specifieke talen (Nederlands en Engels) of door een verschil tussen de status van deze talen (de T1 en de T2). Een visuele-wereld-paradima werd gebruikt voor dit onderzoek. Hierin worden oogbewegingen gemeten terwijl deelnemers luisteren naar zinnen en kijken naar plaatjes op een scherm. Oogbewegingen tijdens het luisteren naar zinnen in de T1 en de T2 met een neutraal of voorspellend werkwoord werden vergeleken (e.g. *Marie leest een brief* versus *Marie steelt een brief* waarbij alle afgebeelde objecten op het scherm 'steelbaar' waren maar enkel één object 'leesbaar'). Nederlands-Engels tweetaligen keken vaker naar de afbeelding van het laatste woord in de zin in the voorspellende dan in de neutrale conditie in beide talen, nog voordat het horen van dat woord oogbewegingen kon beïnvloeden. Eentaligen deden hetzelfde in het Engels. Deze bevinding toont aan dat de eentaligen en de tweetaligen in beide talen op basis van het werkwoord een voorspelling maakten van het object dat nog genoemd zou worden. De mate waarin voorspeld werd verschilde niet tussen de T1 en T2 binnen tweetaligen, en ook niet tussen de T1 van eentaligen en de T2 van tweetaligen.

In dit paradigma zouden de voorspellingen gebaseerd kunnen zijn op overkoepelende betekenis van de zin (d.w.z. op basis van de betekenis van *Marie breit*). Anderzijds is het goed mogelijk dat een mechanisme op basis van eenvoudige associaties tussen woordbetekenisssen betrokken was, waarbij er door activatie van de betekenis van '*breit*' activatie uitspreid naar gerelateerde items zoals *'sjaal*'. De bevindingen zijn dan ook in overeenstemming met de theorie van Pickering and Gambi (2018) omtrent voorspelling tijdens taalverwerking. Deze theorie gaat uit van twee routes tot voorspellen: voorspellen-via-taalproductie en voorspellen-via-associaties. De eerste route is het meest effectief, maar deze is ook afhankelijk van de beschikbare tijd en cognitieve middelen (zoals werkgeheugen). De tweede route verloopt automatisch en is dus niet optioneel.

Aangezien voorspellen-via-associaties een automatisch proces is, zou deze route grotendeels onaangedaan moeten zijn in populaties met minder werkgeheugencapaciteit, zoals T2-verwerkers.

In **HOOFDSTUK 3** onderzochten we voorspelling op het semantische niveau op een meer fijnmazige wijze, opnieuw met het visuele-wereld-paradigma. De hypothese was dat subtiele verschillen tussen voorspelling in T1 en T2 zouden ontstaan wanneer de omstandigheden meer cognitieve middelen vereisten, die misschien niet beschikbaar zijn tijdens T2-verwerking. De gebruikte zinnen in het experiment waren meer gevarieerd, langer en hadden vaak een meer complexe zinsbouw dan de zinnen gebruikt in **HOOFDSTUK 2**. Het doelwoord aan het einde van de zin kon voorspeld worden op basis van de betekenis de alle woorden in de zin samen. De afbeeldingen kwamen bovendien maar kort voor het te voorspellen woord gezegd werd op het scherm te staan, om woordactivatie als gevolg van het zien van het plaatje zoveel mogelijk te beperken. We onderzochten ook of het uitspreiden van semantische activatie verschilde tussen T1 en T2 door een conditie toe te voegen waarin een afbeelding van een semantische concurrent van het te voorspellen woord (bijvoorbeeld *arm* voor doelwoord *been*) op het scherm stond. In overeenstemming met de bevindingen in **HOOFDSTUK 2**, vonden we ook hier dat tweetaligen zowel in T1 als in T2 voorspellingen maken op basis van semantische informatie uit de zin. De Engels-Nederlands tweetalige deelnemers keken namelijk meer naar de afbeelding van het doelwoord dan naar ongerelateerde woorden in beide talen. In dit geval vonden we echter wél een effect van T2: zowel pre-activatie van het doelwoord als van de semantische concurrent was zwakker in deze taal. De pre-activatie van de concurrent was bovendien niet alleen afhankelijk van de taal, maar ook van de sterkte van de relatie tussen het doelwoord en de concurrent. Het effect van de sterkte van de semantische relatie was zwakker en begon later in de T2 dan in de T1. In meer uitdagende omstandigheden blijkt er dus weldegelijk een nadeel te zijn van T2 bij het maken van semantischische voorspellingen. Tragere of zwakkere spreiding van activatie in de T2 zou het gevolg kunnen zijn van zwakkere verbindingen tussen woordvormen en hun semantiek in T2 als gevolg van het minder gebruiken van die taal (e.g. Gollan et al., 2008).

Verschillen tussen semantische voorspelling in T1 en T2 worden niet consequent gevonden, en het is onduidelijk welke mechanismen aan het verschil ten grondslag liggen, wanneer het verschil wél aanwezig is. In **HOOFDSTUK 4** richtten we ons daarom op de factoren die mogelijk aan het verschil ten grondslag liggen. De eerste factor is de cognitieve belasting die T2-verwerking mogelijk met zich meebrengt (e.g., Francis & Gutiérrez, 2012; McDonald, 2006). Pickering and Gambi (2018) stelden voor dat voorspelling-via-taalproductie zwakker is bij mensen met minder cognitieve middelen, aangezien deze route tijd en cognitieve middelen vereist. Indien cognitieve belasting inderdaad ten grondslag ligt aan het zwakker voorspellen in T2, dan zou het extra verhogen van de belasting met name het voorspellen in T2 moeten treffen. De tweede mogelijke onderliggende factor, verwerkingssnelheid, werd gekozen op basis van een vergelijkbare redenering. T2-verwerking is doorgaans trager dan T1-verwerking (Moreno et al., 2008; Shook, Goldrick, Engstler, & Marian, 2015). Indien voorspelling-via-productie tijd vereist zou de vertraging in T2-verwerking de oorzaak kunnen zijn van het verschil tussen voorspellen in T1 en T2. Vertraagd presenteren van zinnen in T2 zou het nadelige effect in T2 dan moeten verlichten. In twee experimenten manipuleerden we cognitieve belasting (door deelnemers te vragen tijdens de helft van de te beluisteren zinnen 9 lettergrepen te onthouden) en verwerkingssnelheid (door de helft van de zinnen in de T1 te versnellen en de helft in de T2 te vertragen. Dezelfde materialen als in **HOOFDSTUK 3** werden gebruikt, zodat ook kon worden bekeken of we de resultaten van **HOOFDSTUK 4** konden repliceren.

In beide experimenten repliceerden we de bevinding dat het voorspellen van doelwoorden zwakker was in de T2, in de conditie zonder toegevoegde cognitieve belasting (Experiment 1) of snelheidsmanipulatie (Experiment 2). Ook pre-activatie van semantische concurrenten was zwakker in de T2 dan in de T1 (zoals in **HOOFDSTUK 3**), maar enkel in de volledige dataset (dus inclusief de data in de conditie met toegevoegde cognitieve belasting en snelheidsmanipulatie). De bevinding in **HOOFDSTUK 3,** dat pre-activatie van de semantische concurrent later en/of zwakker werd beïnvloed door de sterkte van de semantische relatie tussen doelwoord en concurrent in T2, werd echter niet gerepliceerd

In Experiment 1 in **HOOFDSTUK 4** vonden we dat tweetalige deelnemers zwakker voorspellen als hun werkgeheugen extra belast werd. Dit effect van belasting was in beide talen aanwezig, maar tegen de verwachtingen in was het effect groter in T1. In Experiment 2 (**HOOFDSTUK 4**) werd minder voorspelling van doelwoorden gevonden tijdens versnelde zinnen in T1, maar vertraging had geen versterkend effect op voorspellen in T2. De experimenten bevestigen dus dat verwerkingssnelheid en werkgeheugen betrokken zijn bij het voorspellen tijdens taalverwerking (Huettig & Janse, 2016; Ito et al., 2017). Cognitieve belasting en verwerkingssnelheid hadden met name een effect op het voorspellen van doelwoorden in T1. Mogelijk beïnvloedt T2 het voorspellen dus niet direct, maar heeft het invloed op het gewicht dat toegekend wordt aan verschillende routes tot voorspellen. In T1 gebruiken tweetaligen misschien meer dan in T2 de route tot voorspellen die werkgeheugen vereist. Het voorspellen van semantische concurrenten werd in geen van beide talen beïnvloedt door cognitieve belasting, hetgeen suggereert dat hiervoor met name gebruikt werd gemaakt van de route die geen cognitieve middelen vereist (bijv. voorspellen-via-associaties). Het versnellen van de zinnen had geen effect op het voorspellen van semantische concurrenten in T1, maar in T2 werd het voorspellen wel versterkt door tragere presentatie. De bevinding dat cognitieve belasting en verwerkingssnelheid het voorspellen beïnvloedde in de T1 en de T2 is opnieuw in overeenstemming met het idee dat dezelfde mechanismen aan voorspellen ten grondslag liggen in T1 en T2 (Kaan, 2014). Ook komen de resultaten overeen met bevindingen uit eerder onderzoek waaruit bleek dat cognitieve middelen en verwerkingssnelheid een rol spelen in het voorspellen tijdens taalverwerking (Huettig & Janse, 2016; Ito et al., 2017).

Een aantal prominente theoriën over voorspellen tijdens taalverwerking nemen aan dat linguïstische representaties gebruikt voor taalbegrip en taalproductie dezelfde zijn (e.g., Pickering & Gambi, 2018; Pickering & Garrod, 2013). Dit brengt met zich mee dat het luisteren naar spraak mogelijk invloed heeft op spraakproductie daarna. In **HOOFDSTUK 5** hebben we onderzocht of dat ook gebeurt in T2. Nederlands-Engels tweetalige deelnemers lazen voor en na het luisteren naar een T1-spreker van het Engels, hardop Engelse zinnen voor met doelfonemen /æ/ en /b/ in woordfinale positie. De tweetaligen pasten hun uitspraak van doelfoneem /b/ aan na het luisteren naar de T1-spreker, maar de uitspraak van /æ/ werd niet aangepast aan de T1-spreker. De mate waarin de deelnemers hun uitspraak aanpasten werd niet beïnvloed door de fysieke aanwezigheid van de T1-spreker van het Engels. Ook het aantal zinnen tussen de productie van het foneem door de T1-spreker en de productie van het foneem door de deelnemer, had geen effect op de mate van aanpassing. Indien spraakaanpassing een volledig automatisch proces was, dan zou niet alleen de deelnemer, maar ook de T1-spreker aanpassing van uitspraak moeten laten zien na het luisteren naar één of meerdere deelnemers. Er werd echter geen consequente aanpassing van de uitspraak van de T1-spreker gevonden. Deze bevindingen laten zien dat tweetaligen inderdaad hun uitspraak van T2-fonemen kunnen aanpassen na het luisteren naar een T1-spreker van die taal, maar dat de uitspraak niet voor elke klank wordt aangepast.

De studies gerapporteerd in dit proefschrift dragen bij aan het onderzoeksveld van voorspellen tijdens taalverwerking, en dat van gesproken taalverwerking door tweetaligen. Ten eerste toonden we aan dat tweetaligen semantische informatie kunnen voorspellen op basis van informatie uit de zinscontext, zelfs tot op hetzelfde niveau als in de T1. In een moeilijkere context was het voorspellen zwakker in T2 dan in T1, waarbij we de talen vergeleken binnen dezelfde deelnemers.

Ten tweede vonden we dat voorspellen via spreidende activatie trager verloopt in T2, hetgeen mogelijk verklaard kan worden door zwakkere verbindingen tussen woordvorm en semantiek. Verder vonden we geen bewijs dat cognitieve belasting en verwerkingssnelheid ten grondslag liggen aan het zwakker voorspellen in T2, maar deze factoren zouden samen met eventuele andere factoren wel kunnen zorgen voor een verschuiving in voorspelstrategie. Hierbij zouden tweetaligen in T2 minder gebruik maken van voorspellen-via-productie, dat cognitieve middelen vereist, dan in T1. De factoren belasting en verwerkingssnelheid hadden wel invloed op het voorspellen in beide talen, hetgeen in overeenstemming is met de notie dat de mechansimen die aan voorspellen tijdens taalverwerking ten grondslag liggen dezelfde zijn in T1 en T2. Ten laatste, luisteren de T2, gesproken door een T1-spreker, kan leiden tot veranderingen in spraakproductie in T2-sprekers, maar niet voor elk foneem.

REFERENCES

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 1029–1040. https://doi.org/10.1037/a0015901

Dell, G. S., & Chang, F. (2013). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120394–20120394. https://doi.org/10.1098/rstb.2012.0394

European Commission. (2012). Europeans and their languages. Retrieved from http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_386_en.pdf

Foucart, A., Ruiz-Tada, E., & Costa, A. (2015). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition*, 1–7. https://doi.org/10.1017/S1366728915000486

Francis, W. S., & Gutiérrez, M. (2012). Bilingual recognition memory: Stronger performance but weaker levels-of-processing effects in the less fluent language. *Memory & Cognition*, *40*(3), 496–503. https://doi.org/10.3758/s13421-011-0163-3

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209. https://doi.org/10.1037/a0022256

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, *30*(3), 251–266. https://doi.org/10.1023/A:1010490917575

Hopp, H. (2015). Semantics and Morphosyntax in Predictive L2 Sentence Processing. *International Review of Applied Linguistics.* https://doi.org/10.1515/iral-2015-0014

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459

Ito, A., Corley, M., & Pickering, M. J. (2017). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, *21*(2), 251–264. https://doi.org/10.1017/S1366728917000050

Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, *98*, 1–11. https://doi.org/10.1016/j.jml.2017.09.002

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, *4*(2), 257–282. https://doi.org/10.1075/lab.4.2.05kaa

Kaan, E., Kirkham, J., & Wijnen, F. (2014). Prediction and integration in native and second-language processing of elliptical structures. *Bilingualism: Language and Cognition*, 1–18. https://doi.org/10.1017/S1366728914000844

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 190–207).

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*(3), 381–401. https://doi.org/10.1016/j.jml.2006.06.006

Moreno, E. M., Rodríguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics*, *21*(6), 477–508. https://doi.org/10.1016/j.jneuroling.2008.01.003

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: a theory and review. *Psychological Bulletin*. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(04), 329–347. https://doi.org/10.1017/S0140525X12001495

Schmidtke, J. (2016). The Bilingual Disadvantage in Speech Understanding in Noise Is Likely a Frequency Effect Related to Reduced Language Exposure. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00678

Shook, A., Goldrick, M., Engstler, C., & Marian, V. (2015). Bilinguals Show Weaker Lexical Access During Spoken Sentence Comprehension. *Journal of Psycholinguistic Research*, *44*(6), 789–802. https://doi.org/10.1007/s10936-014-9322-6

Van Berkum, J. J. A. (2010). The brain is a prediction machine that cares about good and bad—any implications for neuropragmatics. *Italian Journal of Linguistics*, *22*(1), 181–208.

Weber, A., & Broersma, M. (2012). Spoken word recognition in second language acquisition. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Hoboken, NJ, USA: John Wiley & Sons, Inc. Retrieved from http://doi.wiley.com/10.1002/9781405198431

APPENDICES CHAPTER 2

APPENDIX 2A

Stimuli sets in English and Dutch.

Set	Constraining English (Dutch)	Neutral English (Dutch)	Target object English (Dutch)	Distractor 1 English (Dutch)	Distractor 2 English (Dutch)	Distractor 3 English (Dutch)
1	cuts (snijdt)	shares (deelt)	cheese (kaas)	mailbox (brievenbus)	desk (bureau)	bottle (fles)
2	reads (leest)	steals (pikt)	letter (brief)	car (auto)	backpack (rugzak)	wheelchair (rolstoel)
3	tastes (proeft)	grabs (grijpt)	carrot (wortel)	dress (jurk)	napkin (zakdoek)	ruler (lat)
4	climbs (beklimt)	fixes (repareert)	wall (muur)	lock (slot)	heel (hak)	shower (douche)
5	boils (kookt)	buys (koopt)	potato (aardappel)	coat (jas)	flower (bloem)	dog (hond)
6	visits (bezoekt)	prefers (verkiest)	church (kerk)	skirt (rok)	sausage (worst)	belt (riem)
7	closes (sluit)	reveals (onthult)	window (raam)	horse (paard)	needle (naald)	spoon (lepel)
8	trains (traint)	catches (vangt)	horse (paard)	letter (brief)	plate (bord)	sword (zwaard)
9	folds (vouwt)	saves (bewaart)	dress (jurk)	bucket (emmer)	candle (kaars)	chest (kist)
10	knits (breit)	loses (verliest)	scarf (sjaal)	cheese (kaas)	comb (kam)	barrel (ton)
11	sharpens (slijpt)	measures (meet)	pencil (potlood)	church (kerk)	rope (touw)	tree (boom)
12	builds (bouwt)	breaks (breekt)	fence (hek)	pencil (potlood)	necklace (ketting)	mirror (spiegel)
13	empties (leegt)	paints (verft)	mailbox (brievenbus)	window (raam)	roof (dak)	chair (stoel)
14	irons (strijkt)	designs (ontwerpt)	skirt (rok)	wall (muur)	bicycle (fiets)	kite (vlieger)
15	drives (bestuurt)	takes (neemt)	car (auto)	potato (aardappel)	knife (mes)	shovel (schop)
16	wears (draagt)	chooses (kiest)	coat (jas)	fence (hek)	mountain (berg)	bridge (brug)
17	opens (opent)	draws (tekent)	lock (slot)	carrot (wortel)	wig (pruik)	raft (vlot)
18	fills (vult)	throws (gooit)	bucket (emmer)	scarf (sjaal)	key (sleutel)	whistle (fluit)

Appendices Chapter 3

Appendix 3A

Sentences with targets and competitors

Stimulus set	English sentence	Dutch sentence	Target English	Target Dutch	Competitor English	Competitor Dutch	Association strength L2	Association strength L1
1	The man went sailing on his	De man ging zeilen op zijn	boat	boot	anchor	anker	0.57	0.49
1	The sailor had a tattoo depicting an	De zeeman had een tattoo van een	anchor	anker	boat	boot	0.57	0.49
2	Eric had beautiful guppies and a turtle in his	Erik had prachtige guppy's en een schildpad in zijn	aquarium	aquarium	shark	haai	0.60	0.48
2	Surfers are scared of getting bitten by a	Surfers zijn bang om gebeten te worden door een	shark	haai	aquarium	aquarium	0.60	0.48
3	An insect crawled over her	Een insect kroop over haar	arm	arm	leg	been	0.25	0.33
3	During his last skiing trip he broke his	Tijdens zijn laatste skireisje brak hij zijn	leg	been	arm	arm	0.25	0.33
4	The policeman collected all the documents and put them in a	De politie verzamelde alle documenten en stopte ze in een	folder	map	backpack	rugzak	0.72	0.70
4	The hiker put his water bottle in his	De wandelaar stopte zijn waterfles in zijn	backpack	rugzak	folder	map	0.72	0.70
5	Santa Claus put a present in his	De Kerstman stopte een cadeautje in zijn	bag	zak	wallet	portefeuille	0.36	0.57
5	He took a euro out of his leather	Hij nam een euro uit zijn leren	wallet	portefeuille	bag	zak	0.36	0.57
6	The monkey peeled a	De aap pelde een	banana	banaan	pineapple	ananas	0.55	0.47
6	The Hawaiian pizza was topped with slices of ham and	De pizza Hawaï was belegd met plakjes ham en	pineapple	ananas	banana	banaan	0.55	0.47
7	He always sang in the	Hij zong altijd onder de	shower	douche	bath	bad	0.36	0.38
7	To relax her muscles she took a	Om haar spieren te ontspannen nam ze een	bath	bad	shower	douche	0.36	0.38
8	He rested his head on a	Hij liet zijn hoofd rusten op een	pillow	kussen	bed	bed	0.37	0.40
8	He was tired so went to	Hij was moe dus ging hij naar	bed	bed	pillow	kussen	0.37	0.40
9	He drove to the garage for a new	Hij reed naar de garage voor een nieuwe	car	auto	bike	fiets	0.46	0.51
9	He rode to school on a	hij reed naar school op de	bike	fiets	car	auto	0.46	0.51
10	The biologist studied the cells through a	De bioloog bestudeerde de cellen door een	microscope	microscoop	binoculars	verrekijker	0.62	0.60
10	He studied the rare bird through his	Hij bestudeerde de zeldzame vogel door zijn	binoculars	verrekijker	microscope	microscoop	0.62	0.60
11	The dog barked at a	De hond blafte naar een	cat	kat	bird	vogel	0.61	0.47
11	She heard the sound of a	Ze hoorde het geluid van een	bird	vogel	cat	kat	0.61	0.47
12	The equipment was sent to the planet in a	De apparatuur werd naar de planeet gestuurd in een	rocket	raket	bomb	bom	0.62	0.52
12	In his backpack the terrorist had a	In zijn rugzak had de terrorist een	bomb	bom	rocket	raket	0.62	0.52
13	Her baby doesn't like drinking from a	Haar baby drinkt niet graag uit een	bottle	fles	glass	glas	0.41	0.37

13	He poured some lemonade into a	Hij schonk wat limonade in een	glass	glas	bottle	fles	0.41	0.37
14	Derrick collects magnets to put on his	Derrick verzamelt magneten voor op zijn	fridge	koelkast	bottle	fles	0.56	0.56
14	He poured a glass of wine and put the cork back in the	Hij schonk een glas wijn in en stopte de kurk terug in de	bottle	fles	fridge	koelkast	0.56	0.56
15	The bird sat on a broken	De vogel zat op een gebroken	branch	tak	pickaxe	houweel	0.78	0.80
15	Coal was extracted with a shovel and a	Steenkool werd gewonnen met een schop en een	pickaxe	houweel	branch	tak	0.78	0.80
16	The janitor cleaned the floor with his bucket and his	De conciërge boende de vloer met zijn emmer en zijn	mop	dweil	brush	borstel	0.60	0.48
16	She sat on her knees and scrubbed the floor with a	Ze zat op haar knieën en schrobde de vloer met een	brush	borstel	mop	dweil	0.60	0.48
17	He ran to the station but missed the	Hij rende naar het station maar miste de	train	trein	bus	bus	0.41	0.38
17	In the USA children are brought to school by	In VS worden kinderen naar school gebracht met een	bus	bus	train	trein	0.41	0.38
18	The Arab rode into the desert on a	De Arabier reed de woestijn in op een	camel	kameel	donkey	ezel	0.57	0.45
18	To bring the goods down from the mountain. he put them on the back of a	Om de goederen de berg af te brengen legde hij ze op de rug van een	donkey	ezel	camel	kameel	0.57	0.45
19	The floor in the Persian Palace was covered with a	De vloer in het Perzische paleis was bedekt met een	carpet	tapijt	chair	stoel	0.69	0.63
19	He came in and threw his bag on a	Hij kwam binnen en gooide zijn tas op een	chair	stoel	carpet	tapijt	0.69	0.63
20	The romantic boy was very old-fashioned. He collected songs and recorded them for her on a	De romantische jongen was erg ouderwets. Hij verzamelde liedjes en nam die voor haar op op een	cassette	cassette	radio	radio	0.62	0.75
20	We listened to the morning news on the	We luisterden naar het ochtendnieuws op de	radio	radio	cassette	cassette	0.62	0.75
21	The mouse ate the	De muis at de	cheese	kaas	sandwich	boterham	0.43	0.49
21	In her lunchbox Mary found fruit and a	In haar lunchtrommel vond Marie fruit en een	sandwich	boterham	cheese	kaas	0.43	0.49
22	The nun listened to the sermon in the	De non luisterde naar de preek in de	church	kerk	priest	priester	0.47	0.32
22	He was baptized by a	Hij werd gedoopt door een	priest	priester	church	kerk	0.47	0.32
23	He wanted to marry her. so he gave her a	Hij wilde met haar trouwen dus gaf hij haar een	ring	ring	clock	klok	0.77	0.77
23	Rob was in a hurry and kept watching the	Rob had haast en bleef maar kijken naar de	clock	klok	ring	ring	0.77	0.77
24	It doesn't matter whether you seal a wine bottle with a cap or a	Het maakt niet uit of je een wijnfles afsluit met een dop of een	cork	kurk	grapes	druiven	0.74	0.60
24	Wine is made of	Wijn wordt gemaakt van	grapes	druiven	cork	kurk	0.74	0.60
25	He cut her hair with the	Hij knipte haar haar met de	scissors	schaar	corkscrew	kurkentrekker	0.73	0.63
25	He opened the wine bottle with a	Hij opende de wijnfles met een	corkscrew	kurkentrekker	scissors	schaar	0.73	0.63
26	The magician pulled the rabbit out of his	De goochelaar trok een konijn uit zijn	hat	hoed	cowboy	cowboy	0.56	0.53
26	His granddad told him stories about an indian and a	Zijn opa vertelde hem een verhaal over een indiaan en een	cowboy	cowboy	hat	hoed	0.56	0.53
27	The guppy was eaten by a large	De guppy werd opgegeten door een grote	fish	vis	shell	schelp	0.73	0.60
27	The mussel closed its	De mossel sloot zijn	shell	schelp	fish	vis	0.73	0.60

28	He heard someone knocking. so he opened the	Hij hoorde iemand kloppen dus hij opende de	door	deur	cupboard	kast	0.59	0.53
28	He put the clean plates back in the	Hij zette de schone borden terug in de	cupboard	kast	door	deur	0.59	0.53
29	The doctor listened to his heart with a	De dokter luisterde naar zijn hart met een	stethoscope	stethoscoop	dentist	tandarts	0.73	0.64
29	He had a painful molar so he went to see a	Hij had een pijnlijke kies dus hij ging naar een	dentist	tandarts	stethoscope	stethoscoop	0.73	0.64
30	It is a nice ring with a small	Het is een mooie ring met een kleine	diamond	diamant	necklace	ketting	0.45	0.64
30	She put the ring on her finger and the bracelet around her wrist. Around her neck she wore a	Ze deed de ring om haar vinger en een armband om haar pols. Om haar nek droeg zij een	necklace	ketting	diamond	diamant	0.45	0.64
31	The young mother bought a new brand of diapers for her	De jonge moeder kocht een nieuw merk luiers voor haar	baby	baby	doctor	dokter	0.60	0.61
31	I wish my daughter had married a lawyer or a	Ik wou dat mijn dochter getrouwd was met een advocaat of een	doctor	dokter	baby	baby	0.60	0.61
32	He put a carrot in the cage of his	Hij legde een wortel in het hok van zijn	rabbit	konijn	dog	hond	0.55	0.57
32	Lola would adopt a cat rather than a	Lola adopteert liever een kat dan een	dog	hond	rabbit	konijn	0.55	0.57
33	A Scottish kilt is a kind of	Een Schotse kilt is een soort	skirt	rok	dress	jurk	0.39	0.39
33	At the prom she wore a blue	Op het gala droeg zij een blauwe	dress	jurk	skirt	rok	0.39	0.39
34	He had a bad cold so he blew his	Hij was erg verkouden dus hij snoot zijn	nose	neus	ear	oor	0.51	0.56
34	She whispered something in his	Ze fluisterde iets in zijn	ear	oor	nose	neus	0.51	0.56
35	The circus owned a tiger and a huge grey	Het circus had een tijger en een enorme grijze	elephant	olifant	rhino	neushoorn	0.50	0.40
35	Ivory is derived from an elephant or a	Ivoor is afkomstig van een olifant of een	rhino	neushoorn	elephant	olifant	0.50	0.40
36	The goods were transported in a	De goederen werden vervoerd in een	truck	vrachtwagen	factory	fabriek	0.67	0.75
36	The clothing was made in a large	De kleding werd gemaakt in een grote	factory	fabriek	truck	vrachtwagen	0.67	0.75
37	The king wore his golden	De koning droeg zijn gouden	crown	kroon	neck	hals	0.71	0.72
37	She wore a colorful scarf around her	Ze droeg een kleurrijke sjaal om haar	neck	hals	crown	kroon	0.71	0.72
38	The natives danced around the	De inboorlingen dansten rond het	fire	vuur	smoke	rook	0.50	0.49
38	The chimney was clogged. so the house was full of	De schoorsteen zat verstopt dus het huis stond vol	smoke	rook	fire	vuur	0.50	0.49
39	The cat was saved from the tree by a	De kat werd uit de boom gered door een	fireman	brandweerman	ladder	ladder	0.72	0.66
39	He was cleaning the windows of the upper floor on a	Hij waste de ruiten van de bovenverdieping op een	ladder	ladder	fireman	brandweerman	0.72	0.66
40	At the villa. he wanted to go swimming in a	Hij wilde bij de villa gaan zwemmen in een	pool	zwembad	fountain	fontein	0.60	0.62
40	He threw a penny into the	Ze gooide een muntje in de	fountain	fontein	pool	zwembad	0.60	0.62
41	He is as clever as a	Hij is zo sluw als een	fox	vos	deer	hert	0.57	0.76
41	He took his gun and shot a	Hij nam zijn geweer en schoot een	deer	hert	fox	vos	0.57	0.76
42	Besides cheese of cow's milk the farmer often makes cheese from the milk of his	Naast kaas van koeienmelk maakt de boer vaak kaas van de melk van zijn	goats	geiten	pig	varken	0.45	0.51

42	Spanish ham is meat from a special kind of	Spaanse ham is vlees van een speciaal soort	pig	varken	goats	geiten	0.45	0.51
43	The rock star put new strings on his	De rockster zette nieuwe snaren op zijn	guitar	gitaar	piano	piano	0.37	0.40
43	With such long fingers. you must play the	Met zulke lange vingers speelt u vast	piano	piano	guitar	gitaar	0.37	0.40
44	The mobster played Russian roulette with his	Het maffialid speelde Russische roulette met zijn	gun	geweer	knife	mes	0.47	0.48
44	He cut his food with a	Hij sneed zijn eten met een	knife	mes	gun	geweer	0.47	0.48
45	The hungry woman ordered a coke. fries. and a	De hongerige vrouw bestelde een cola. friet en een	hamburger	hamburger	icecream	ijsje	0.59	0.60
45	It was a warm day so the spoiled child wanted an	Het was een warme dag dus het verwende kind wilde een	icecream	ijsje	hamburger	hamburger	0.59	0.60
46	The boy dressed up as a train conductor and wore a whistle and a	De jongen verkleedde zich als conducteur en droeg een fluitje en een	hat	pet	basket	mand	0.67	0.77
46	She put the food for the picnic in a	Ze deed het eten voor de picknick in een	basket	mand	hat	pet	0.67	0.77
47	The farmer milked a	De boer melkte een	cow	koe	hay	hooi	0.72	0.68
47	The stable boy took a bale of	De stalknecht nam een baal	hay	hooi	cow	koe	0.72	0.68
48	The doctor held the stethoscope against his	De dokter hield de stethoscoop tegen zijn	chest	borst	heart	hart	0.52	0.50
48	He was in love with her. so he gave her a box of chocolates in the shape of a	Hij was verliefd op haar dus hij gaf haar een doos bonbons in de vorm van een	heart	hart	chest	borst	0.52	0.50
49	Santa Claus travels to the North Pole on a	De Kerstman reist naar de Noordpool op een	sled	slee	rope	touw	0.68	0.64
49	The pirate tied the prisoner's hands with a	De piraat bond de handen van zijn gevangene vast met een	rope	touw	sled	slee	0.68	0.64
50	He buried his head in the sand like an	Hij stak zijn kop in het zand als een	ostrich	struisvogel	kangaroo	kangoeroe	0.63	0.48
50	When he was in Australia. he saw a young joey in the pouch of a	Toen hij in Australië was zag hij een jong in de buidel van een	kangaroo	kangoeroe	ostrich	struisvogel	0.63	0.48
51	The politician kept the secret document and the money in a	De politicus bewaarde het geheime document en het geld in een	safe	kluis	key	sleutel	0.62	0.46
51	He quickly opened the lock with his	Hij opende vlug het slot met zijn	key	sleutel	safe	kluis	0.62	0.46
52	The knight saw his enemy and drew his	De ridder zag zijn vijand en trok zijn	sword	zwaard	king	koning	0.65	0.54
52	He was the prince and his father was	Hij was de prins en zijn vader was een	king	koning	sword	zwaard	0.65	0.54
53	The dragon was slain by the courageous	De draak werd gedood door de dappere	knight	ridder	wizard	tovenaar	0.66	0.62
53	The head of the school of magic was a	Het hoofd van de toverschool was een	wizard	tovenaar	knight	ridder	0.66	0.62
54	He hated the sour taste of	Hij haatte de zure smaak van	lemon	citroen	orange	sinaasappel	0.50	0.39
54	She squeezed the delicious fresh juice from the	Ze perste het heerlijke verse sap uit de	orange	sinaasappel	lemon	citroen	0.50	0.39
55	It is so dark I can barely read. I would like a better	Het is hier zo donker dat ik bijna niet kan lezen. Ik wil graag een betere	light	lamp	candle	kaars	0.49	0.58
55	In church we saw the flickering light of a	In de kerk zag hij het flikkerende licht van een	candle	kaars	light	lamp	0.49	0.58
56	The circus performer tamed a	De circusartiest temde een	lion	leeuw	dragon	draak	0.51	0.55
56	He heard that the beast had two heads and breathed fire. It must have been a	Hij hoorde dat het beest twee koppen had en vuur spuwde. Het was zeker een	dragon	draak	lion	leeuw	0.51	0.55

57	Alexandra put her new clothes on a shelf in her	Alexandra legde haar nieuwe kleding op een plank in haar	closet	kast	lock	slot	0.64	0.57
57	She locked her bicycle to a fence with a	Zij zette haar fiets vast aan een hek met een	lock	slot	closet	kast	0.64	0.57
58	The thief was caught and had to go to	De dief werd gepakt en moest naar de	jail	gevangenis	man	man	0.58	0.60
58	She fell in love with a handsome	Ze werd verliefd op een knappe	man	man	jail	gevangenis	0.58	0.60
59	When you drive. you keep your eyes on the	Als je rijdt houd je je ogen op de	road	weg	map	kaart	0.71	0.63
59	Could you show me where the village is on a	Kun je me laten zien waar het dorpje ligt op een	map	kaart	road	weg	0.71	0.63
60	I saw myself in the	Ik zag mezelf in de	mirror	spiegel	eyes	ogen	0.58	0.49
60	Without her sunglasses. the sun hurt Erika's	Zonder haar zonnebril deed de zon Erika pijn aan haar	eyes	ogen	mirror	spiegel	0.58	0.49
61	The adventurer started to climb a	De avonturier begon aan de beklimming van een	mountain	berg	rock	steen	0.65	0.68
61	The little frog sat on a	Het kleine kikkertje zat op een	rock	steen	mountain	berg	0.65	0.68
62	The cat killed a	De kat doodde een	mouse	muis	cage	kooi	0.64	0.60
62	He petted his parrot and then put it back in its	Hij aaide zijn papegaai en zette hem toen terug in zijn	cage	kooi	mouse	muis	0.64	0.60
63	The strongest finger on your hand is your	De sterkste vinger aan je hand is je	thumb	duim	stamp	postzegel	0.77	0.80
63	He mailed the letter without a	Hij verstuurde de brief zonder een	stamp	postzegel	thumb	duim	0.77	0.80
64	The chicken laid an	De kip legde een	egg	ei	nest	nest	0.54	0.59
64	In spring the birds built a	In het voorjaar bouwden de vogels een	nest	nest	egg	ei	0.54	0.59
65	The little boy marched like a	Het kleine jongetje marcheerde als een	soldier	soldaat	nurse	verpleegster	0.66	0.71
65	During the war. she worked at the hospital as a	Tijdens de oorlog werkte zij in een ziekenhuis als	nurse	verpleegster	soldier	soldaat	0.66	0.71
66	To protect her fingers from the cold she wore a	Om haar vingers tegen de kou te beschermen droeg ze een	glove	handschoen	package	pakketje	0.75	0.74
66	This morning. someone delivered us a	Vanochtend bezorgde iemand ons een	package	pakketje	glove	handschoen	0.75	0.74
67	Clara put the flowers in an expensive	Clara zette de bloemen in een dure	vase	vaas	painting	schilderij	0.58	0.61
67	The artist took his brush and made a	De kunstenaar pakte zijn kwast en maakte een	painting	schilderij	vase	vaas	0.58	0.61
68	He had a hole in his	Hij had een gat in zijn	pants	broek	sweater	trui	0.49	0.51
68	For Christmas. she knitted her son a	Voor kerst breidde ze voor haar zoon een	sweater	trui	pants	broek	0.49	0.51
69	Dick wrote a chapter in the	Dick schreef een hoofdstuk in het	book	boek	paper	papier	0.60	0.65
69	Jot it down on a piece of	Noteer het op een stuk	paper	papier	book	boek	0.60	0.65
70	The boy at the zoo brought bananas to feed a	De jongen in de dierentuin bracht bananen mee om te voeren aan een	monkey	aap	parrot	papegaai	0.61	0.57
70	The colorful bird that repeats your words is called a	De kleurrijke vogel die je woorden herhaalt heet een	parrot	papegaai	monkey	aap	0.61	0.57
71	The clown sold her father a	De clown verkocht haar vader een	balloon	ballon	plane	vliegtuig	0.66	0.60
71	The pilot entered the cockpit of the	De piloot betrad de cockpit van het	plane	vliegtuig	balloon	ballon	0.66	0.60

72	It was his birthday and his mother baked a	Hij was jarig en zijn moeder bakte een	cake	taart	potato	aardappel	0.71	0.69
72	Max wanted to help his mother in the kitchen. so he peeled a	Max wilde zijn moeder helpen in de keuken. dus schilde hij een	potato	aardappel	cake	taart	0.71	0.69
73	The ceremony was attended by the king and	De ceremonie werd bijgewoond door de koning en	queen	koningin	witch	heks	0.73	0.58
73	She was burned in the middle ages because they thought she was a	Ze werd in de middeleeuwen verbrand want ze hielden haar voor een	witch	heks	queen	koningin	0.73	0.58
74	He worked on a ship as a	Hij werkte op een schip als	sailor	matroos	raft	vlot	0.73	0.70
74	To leave the deserted island. they built a	Om van het onbewoonde eiland af te komen bouwden ze een	raft	vlot	sailor	matroos	0.73	0.70
75	I got sick from eating a poisonous	Ik werd ziek door het eten van een giftige	mushroom	paddenstoel	rose	roos	0.74	0.72
75	She removed the thorns from the red	ze verwijderde de doorns van de rode	rose	roos	mushroom	paddenstoel	0.74	0.72
76	He is so good at horseback riding. He doesn't even use a	Hij is zo goed in paardrijden. Hij gebruikt niet eens een	saddle	zadel	helmet	helm	0.71	0.69
76	He rides a motorbike but he never wears a	Hij rijdt motor maar hij draagt nooit een	helmet	helm	saddle	zadel	0.71	0.69
77	The captain decided to stay with the sinking	De kapitein besloot om te blijven op het zinkende	ship	schip	bridge	brug	0.66	0.50
77	To get to the other side of the river you have to cross a	Om aan de andere kant van de rivier te komen moet je over een	bridge	brug	ship	schip	0.66	0.50
78	The boy enjoyed himself in the pool. He loved going down the	De jongen vermaakte zich in het zwembad. Hij ging graag van de	slide	glijbaan	tripod	statief	0.74	0.64
78	To keep the camera steady. he put it on a	Om de camera recht te houden zette hij hem op een	tripod	statief	slide	glijbaan	0.74	0.64
79	The treasure map was made by a	De schatkaart werd gemaakt door een	pirate	piraat	submarine	duikboot	0.71	0.64
79	To research The Titanic. the research team used a	Om de titanic te bereiken gebruikte het onderzoeksteam een	submarine	duikboot	pirate	piraat	0.71	0.64
80	The player’s cap protected him from the	De pet van de speler beschermde hem tegen de	sun	zon	rain	regen	0.53	0.54
80	He walked outside in the wind and the	Hij liep buiten in de wind en de	rain	regen	sun	zon	0.53	0.54
81	That night he slept at the festival in a	Die nacht sliep hij op het festival in een	tent	tent	house	huis	0.69	0.64
81	He placed a new kitchen in his	Hij plaatste een nieuwe keuken in zijn	house	huis	tent	tent	0.69	0.64
82	He hung the sock on the line with a	Hij hing de sok aan de lijn met een	clothespin	wasknijper	thread	draad	0.68	0.71
82	I sewed on the button with a needle and	Ik naaide de knoop eraan met naald en	thread	draad	clothespin	wasknijper	0.68	0.71
83	The little girl needed to pee. so she went to the	Het kleine meisje moest plassen dus ze ging naar het	toilet	toilet	sink	gootsteen	0.60	0.45
83	She washed the dirty dishes in the	Ze deed de afwas in de	sink	gootsteen	toilet	toilet	0.60	0.45
84	Ron was shocked by the environmental pollution. The whole beach was full of	Ron was geschrokken van de milieuvervuiling. Het hele strand lag vol met	garbage	afval	diaper	luier	0.71	0.79
84	She picked up her baby. It was time to change his	Ze pakte haar baby op. Het was tijd voor het verschonen van zijn	diaper	luier	garbage	afval	0.71	0.79
85	It was raining heavily so Jenny went outside with her	Het regende erg hard dus Jennie ging naar buiten met haar	umbrella	paraplu	torch	fakkel	0.73	0.78
85	To show us the murals in the cave. he lit up a	Om ons de muurschildering in de grot te laten zien ontstak hij een	torch	fakkel	umbrella	paraplu	0.73	0.78
86	He did not want to spill anything so he poured the lemonade through a	Hij wilde niets morsen dus schonk hij de limonade door een	funnel	trechter	volcano	vulkaan	0.78	0.78

86	Lava is the molten rock expelled by a	Lava is gesmolten gesteente dat wordt uitgestoten door een	volcano	vulkaan	funnel	trechter	0.78	0.78
87	The dog looked outside through a	Het hondje keek naar buiten door een	window	raam	roof	dak	0.54	0.53
87	He climbed on top of his house and sat down on the	Hij klom op zijn huis en ging zitten op het	roof	dak	window	raam	0.54	0.53
88	The sommelier handed her the glass and she took a sip of	De sommelier gaf haar het glas en ze nam een slokje	wine	wijn	table	tafel	0.63	0.72
88	He put the chair under a	Hij zette de stoel onder een	table	tafel	wine	wijn	0.63	0.72
89	The dog wagged its	Het hondje kwispelde met zijn	tail	staart	wing	vleugel	0.70	0.74
89	The bird couldn't fly because he had a broken	Het vogeltje kon niet vliegen want hij had een gebroken	wing	vleugel	tail	staart	0.70	0.74
90	The squirrel ate an	De eekhoorn at een	acorn	eikel	tree	boom	0.78	0.67
90	The dog chased our cat up a	De hond joeg onze kat in een	tree	boom	acorn	eikel	0.78	0.67
91	Floris is as slow as a	Floris is zo traag als een	snail	slak	ant	mier	0.72	0.55
91	The insect that can carry fifty times its own weight is called an	Het insect dat vijftig keer zijn eigen gewicht kan dragen heet een	ant	mier	snail	slak	0.72	0.55
92	On Halloween he carved a face out of a	Met Halloween sneed hij een gezicht uit een	pumpkin	pompoen	apple	appel	0.65	0.69
92	Snow White took a bite of her	Sneeuwwitje nam een hap van haar	apple	appel	pumpkin	pompoen	0.65	0.69
93	The Indian carried a bow and an	De indiaan droeg een boog en een	arrow	pijl	needle	naald	0.73	0.78
93	She repaired the skirt with thread and	Ze repareerde de rok met draad en	needle	naald	arrow	pijl	0.73	0.78
94	He tossed the empty plastic cup in a	Hij gooide het lege plastic bekertje in een	trashcan	vuilbak	ashtray	asbak	0.66	0.57
94	He put the cigarette out in the	Ze maakte de sigaret uit in de	ashtray	asbak	trashcan	vuilbak	0.66	0.57
95	The lumberjack chopped wood with his	De houthakker hakte hout met zijn	axe	bijl	hammer	hamer	0.56	0.59
95	He slammed the nail into the wall with a	Hij sloeg de spijker in de muur met een	hammer	hamer	axe	bijl	0.56	0.59
96	The cashier put the groceries into a	De caissière stopte de boodschappen in een	bag	tas	zipper	rits	0.65	0.68
96	This coat has buttons. but I prefer a	Deze jas heeft knopen maar ik verkies een	zipper	rits	bag	tas	0.65	0.68
97	The other player threw the	De andere speler gooide de	ball	bal	racket	tennisracket	0.61	0.66
97	Nadal bought a new	Nadal kocht een nieuw	racket	tennisracket	ball	bal	0.61	0.66
98	He eats out because he is a lousy	Hij gaat uiteten want hij is een slechte	cook	kok	barbecue	barbecue	0.61	0.70
98	He liked to grill meat in summer so he put coals in his	Hij hield ervan 's zomers vlees te grillen en legde kolen in zijn	barbecue	barbecue	cook	kok	0.61	0.70
99	The student repaired his tire and filled it using a	De student plakte zijn band en vulde hem met een	pump	fietspomp	barrel	vat	0.70	0.76
99	Wine is often stored in a wooden	Wijn wordt vaak opgeslagen in een houten	barrel	vat	pump	fietspomp	0.70	0.76
100	He was afraid to catch a cold. so he wore a	Hij was bang verkouden te worden dus hij droeg een	scarf	sjaal	towel	handdoek	0.67	0.68
100	She dried her wet feet with a	Zij droogde haar natte voeten met een	towel	handdoek	scarf	sjaal	0.67	0.68
101	He hit the burglar in the face with a	Hij sloeg de inbreker in het gezicht met een	bat	knuppel	fist	vuist	0.76	0.66

101	He wanted to hit him in the face. so he made a	Hij wilde hem in het gezicht slaan. dus hij maakte een	fist	vuist	bat	knuppel	0.76	0.66
102	The hare will always be faster than the	De haas zal altijd sneller zijn dan de	turtle	schildpad	bat	vleermuis	0.68	0.63
102	High up in the cave they saw a	Hoog boven in de grot zagen ze een	bat	vleermuis	turtle	schildpad	0.68	0.63
103	The colorful bird cracked a nut with its	De gekleurde vogel kraakte een noot met zijn	beak	bek	worm	worm	0.71	0.71
103	The bird ate a big fat	De vogel at een grote dikke	worm	worm	beak	bek	0.71	0.71
104	The boy looked at the long neck of the	De jongen keek naar de lange nek van de	giraffe	giraf	bear	beer	0.64	0.56
104	The child could not sleep without his brown	Het kind kon niet slapen zonder zijn bruine	bear	beer	giraffe	giraf	0.64	0.56
105	He grabbed a razor and shaved his	Hij pakte een scheermes en scheerde zijn	beard	baard	wig	pruik	0.49	0.49
105	She lost her hair so now she wears a	Ze verloor haar haar dus nu draagt ze een	wig	pruik	beard	baard	0.49	0.49
106	The flower was pollinated by a	De bloem werd bestoven door een	bee	bij	girl	meisje	0.77	0.63
106	The boy kissed a	De jongen kuste een	girl	meisje	bee	bij	0.77	0.63
107	The angry driver used his	De boze automobilist gebruikte zijn	horn	claxon	bell	bel	0.70	0.72
107	When it was time to go back to class the students would hear the sound of a	Wanneer het tijd was om terug naar de klas te gaan hoorden de leerlingen het geluid van een	bell	bel	horn	claxon	0.70	0.72
108	The policeman attached him to the fence with	De agent bond hem aan het hek met	handcuffs	handboeien	belt	riem	0.66	0.61
108	To keep up his pants he used a	Om zijn broek op te houden gebruikte hij een	belt	riem	handcuffs	handboeien	0.66	0.61
109	Cinderella scrubbed the	Assepoester boende de	floor	vloer	bench	bank	0.73	0.66
109	The old man in the park sat on a	De oude man in het park ging zitten op een	bench	bank	floor	vloer	0.73	0.66
110	Nikkie hung the colorful painting up on the	Nikkie hing het kleurrijke schilderij aan de	wall	muur	block	blok	0.60	0.70
110	To start building a tower the little boy picked up a wooden	Om te beginnen een toren te bouwen pakte de kleine jongen een houten	block	blok	wall	muur	0.60	0.70
111	Walking through the dark room. I accidentally stubbed my	Rondlopend in het donker stootte ik per ongeluk mijn	toe	teen	feather	veer	0.75	0.70
111	Before there were pens. people wrote with a	Voordat er pennen waren schreef men met een	feather	veer	toe	teen	0.75	0.70
112	The knight took his sword and mounted his	De ridder nam zijn zwaard en besteeg zijn	horse	paard	bow	strik	0.73	0.68
112	He tied the ribbon into a	Hij knoopte het lint in een	bow	strik	horse	paard	0.73	0.68
113	He couldn't see without his	Hij kon niet zien zonder zijn	glasses	bril	bowl	kom	0.70	0.69
113	He poured the soup into a	Hij schonk de soep in een	bowl	kom	glasses	bril	0.70	0.69
114	Bob took all the toys and put them in a	Bob pakte al het speelgoed en deed het in een	box	doos	drawer	lade	0.55	0.61
114	We keep the forks and knives in a	We bewaren de vorken en messen in een	drawer	lade	box	doos	0.55	0.61
115	He already had two girls so this time he hoped for a	Hij had al twee meisjes dus deze keer hoopte hij op een	boy	jongen	woman	vrouw	0.52	0.57
115	He left his wife for another	Hij verliet zijn echtgenote voor een andere	woman	vrouw	boy	jongen	0.52	0.57

116	There was a hole in the sole of the	Er zat een gat in de zool van de	shoe	schoen	bra	beha	0.64	0.65
116	She was a feminist in the sixties and she burned her	Ze was een feministe in de jaren 60 en verbrandde haar	bra	beha	shoe	schoen	0.64	0.65
117	The kids fed the ducks some	De kinderen voerden de eendjes wat	bread	brood	cookie	koekje	0.69	0.64
117	Before going to bed. the boy had milk and a	Voordat hij naar bed ging kreeg het jongetje melk en een	cookie	koekje	bread	brood	0.69	0.64
118	Santa Claus enters your house through the	De Kerstman komt je huis binnen door de	chimney	schoorsteen	bricks	bakstenen	0.72	0.72
118	The house was made of red	Het huis was gemaakt van rode	bricks	bakstenen	chimney	schoorsteen	0.72	0.72
119	The old witch flew off on a	De oude heks vloog weg op een	broom	bezem	wheelbarrow	kruiwagen	0.69	0.68
119	The gardener moved the heavy rocks in a	De tuinman verplaatste de zware stenen in een	wheelbarrow	kruiwagen	broom	bezem	0.69	0.68
120	The lawn was very dry so he watered it with a	Het gazon was erg droog dus hij besproeide het met een	hose	tuinslang	bucket	emmer	0.66	0.56
120	He played in the sand with a shovel and a	Hij speelde op het zand met een schepje en een	bucket	emmer	hose	tuinslang	0.66	0.56
121	He lost his legs so now he has a	Hij verloor zijn benen dus nu heeft hij een	wheelchair	rolstoel	buggy	buggy	0.64	0.61
121	She walked through the zoo with the toddler in a	Ze wandelde door de dierentuin met de peuter in een	buggy	buggy	wheelchair	rolstoel	0.64	0.61
122	The farmer gave them a fresh egg from his	De boer gaf hen een vers ei van zijn	chicken	kippen	butcher	slager	0.64	0.76
122	There were no more lamb chops at the supermarket so I asked the	Er waren geen lamskoteletjes meer in de supermarkt dus ik ging naar de	butcher	slager	chicken	kippen	0.64	0.76
123	He didn't like frying things in oil so he used	Hij hield er niet van dingen te bakken in olie dus gebruikte hij	butter	boter	onion	ui	0.60	0.64
123	Mary's eyes teared up from cutting an	Maries ogen traanden van het snijden van een	onion	ui	butter	boter	0.60	0.64
124	It was raining but the sun was shining. and Maya saw a	Het regende maar de zon scheen en Maya zag een	rainbow	regenboog	butterfly	vlinder	0.72	0.74
124	His last collection included a purple	Tot zijn laatste collectie behoorde een paarse	butterfly	vlinder	rainbow	regenboog	0.72	0.74
125	The jeans closed with a zipper and a	De jeans sloot met een rits en een	button	knoop	suit	pak	0.74	0.72
125	He looked like a penguin in that	Hij zag eruit als een pinguïn in dat	suit	pak	button	knoop	0.74	0.72
126	He wasn't good with plants so he bought a	Hij was niet goed met planten dus kocht hij een	cactus	cactus	dinosaur	dinosaurus	0.77	0.77
126	In the Museum of Natural History he saw an enormous skeleton of a	In het Natuurhistorisch museum zag hij een enorm skelet van een	dinosaur	dinosaurus	cactus	cactus	0.77	0.77
127	You forgot to turn on the flash on your	Je bent vergeten de flits aan te zetten op je	camera	camera	phone	telefoon	0.69	0.63
127	She couldn't leave the house. so she called her daughter on the	Ze kon het huis niet uit en belde haar dochter met de	phone	telefoon	camera	camera	0.69	0.63
128	Ron had several blisters on his	Ron had meerdere blaren op zijn	feet	voeten	can	blik	0.70	0.75
128	She wanted to eat peas so she opened a	Ze wilde erwten eten dus ze opende een	can	blik	feet	voeten	0.70	0.75
129	The pretty girl sat at the bar on a	Het mooie meisje zat aan de bar op een	stool	kruk	cane	stok	0.76	0.59
129	To help him walk better. the man used a	Om beter te kunnen lopen gebruikte de man een	cane	stok	stool	kruk	0.76	0.59
130	The farmer tended to his field on a	De boer bewerkte zijn akker op een	tractor	tractor	caravan	caravan	0.75	0.62
130	He thought it was too cold to sleep in a tent so he went on a trip with a	Hij vond het te koud om te slapen in een tent dus hij ging op reis met een	caravan	caravan	tractor	tractor	0.75	0.62

131	Sleeping beauty pricked her finger on a	Doornroosje prikte haar vinger aan een	spinningwheel	spinnewiel	carousel	draaimolen	0.72	0.60
131	She wanted to sit on the pink horse in the	Ze wilde op het roze paard in de	carousel	draaimolen	spinningwheel	spinnewiel	0.72	0.60
132	That green soup is made of	Die groene soep wordt gemaakt van	peas	erwten	carrot	wortel	0.72	0.64
132	To improve his vision. he ate a	Om zijn zicht te verbeteren at hij een	carrot	wortel	peas	erwten	0.72	0.64
133	The knight lived in a	De ridder woonde in een	castle	kasteel	city	stad	0.73	0.67
133	He loved the countryside. but he lived in the	Hij hield van het platteland. maar hij woonde in de	city	stad	castle	kasteel	0.73	0.67
134	The playground only had a slide and a	De speeltuin had enkel een glijbaan en een	swing	schommel	catapult	katapult	0.74	0.69
134	The naughty boy shot rocks at a cat with a	De stoute jongen schoot stenen naar een poes met een	catapult	katapult	swing	schommel	0.74	0.69
135	On top of the cake she put a nice red	Bovenop de taart legde ze een mooie rode	cherry	kers	strawberries	aardbeien	0.54	0.72
135	She made a delicious jam of	Ze maakte een heerlijke jam van	strawberries	aardbeien	cherry	kers	0.54	0.72
136	The pirate found a treasure of gold coins in a	De piraat vond een schat van gouden munten in een	chest	kist	well	put	0.75	0.72
136	In the middle ages people took water from a	In de middeleeuwen haalden mensen water uit een	well	put	chest	kist	0.75	0.72
137	The Cuban smoked a	De Cubaan rookte een	cigar	sigaar	waiter	ober	0.76	0.73
137	In this restaurant you are served by a friendly	In dit restaurant word je bediend door een aardige	waiter	ober	cigar	sigaar	0.76	0.73
138	I couldn't see his face because he wore a	Ik kon zijn gezicht niet zien want hij droeg een	mask	masker	clown	clown	0.63	0.60
138	For his third birthday. his dad dressed up as a	Voor zijn derde verjaardag verkleedde zijn vader zich als	clown	clown	mask	masker	0.63	0.60
139	I heard the hissing of a venomous	Ik hoorde het gesis van een giftige	snake	slang	cock	haan	0.69	0.72
139	Early in the morning he heard the cock-a-doodle-doo of a	Vroeg in de ochtend hoorde hij het gekukeleku van een	cock	haan	snake	slang	0.69	0.72
140	He made a part in his hair with a	Hij maakte een scheiding in zijn haar met een	comb	kam	handkerchief	zakdoek	0.76	0.78
140	He blew his nose into a	Hij snoot zijn neus in een	handkerchief	zakdoek	comb	kam	0.76	0.78
141	He pretended to be with the mafia. but he was actually a	Hij deed alsof hij bij de maffia hoorde maar hij was eigenlijk een	cop	agent	desk	bureau	0.66	0.47
141	He took his textbook and sat at his	Hij nam zijn tekstboek en ging zitten aan zijn	desk	bureau	cop	agent	0.66	0.47
142	Flour for bread is usually made of	Meel voor brood wordt meestal gemaakt van	wheat	graan	corn	mais	0.39	0.78
142	Tortillas are often made of	Tortilla's worden vaak gemaakt van	corn	mais	wheat	graan	0.39	0.78
143	He checked the time on his	Hij keek hoe laat het was op zijn	watch	horloge	couch	bank	0.67	0.73
143	Martin was very lazy today and watched television on the	Martin was erg lui vandaag en keek tv op de	couch	bank	watch	horloge	0.67	0.73
144	The little girl put her savings in a	Het kleine meisje deed haar spaargeld in een	piggybank	spaarpot	wateringcan	gieter	0.73	0.78
144	He watered the flowers with a	Hij gaf de bloemen water met een	wateringcan	gieter	piggybank	spaarpot	0.73	0.78
145	The pope wore a necklace with a	De paus droeg een ketting met een	cross	kruis	flag	vlag	0.73	0.69
145	To show that he had surrendered. he waved a white	Om te laten zien dat hij zich overgaf zwaaide hij met een witte	flag	vlag	cross	kruis	0.73	0.69

146	He put the ring on her	Hij deed de ring om haar	finger	vinger	hair	haar	0.68	0.70
146	She went to the salon to color her	Ze ging naar de kapper voor een kleurtje in haar	hair	haar	finger	vinger	0.68	0.70
147	The English queen drank tea from a	De Engelse koningin dronk thee uit een	cup	kopje	toaster	broodrooster	0.72	0.70
147	Since the slice of bread was a bit old. he put it in the	Omdat de boterham wat oud was deed hij het in de	toaster	broodrooster	cup	kopje	0.72	0.70
148	It was dark so Simon closed the	Het was donker dus Simon sloot de	curtains	gordijnen	iron	strijkijzer	0.63	0.65
148	His shirt was completely wrinkled. so his mother took out her	Zijn hemd was helemaal gekreukeld dus zijn moeder pakte haar	iron	strijkijzer	curtains	gordijnen	0.63	0.65
149	The little girl played with her	Het kleine meisje speelde met haar	doll	pop	purse	handtas	0.68	0.75
149	She walked up to the mirror and took her lipstick from her	Ze liep naar de spiegel en pakte haar lipstick uit haar	purse	handtas	doll	pop	0.68	0.75
150	He made a hole in the wall for the screw with a	Hij maakte een gat in de muur voor de schroef met een	drill	boor	jack	krik	0.69	0.56
150	To replace the tire. the car was lifted with a	Om de autoband te vervangen werd de auto opgetild met een	jack	krik	drill	boor	0.69	0.56
151	To built up suspense. the circus artist beat the	Om de spanning op te bouwen sloeg the circusartiest op een	drum	trommel	kettle	waterkoker	0.75	0.64
151	She offered him tea and heated up the water in a	Ze bood hem thee aan en verwarmde het water in een	kettle	waterkoker	drum	trommel	0.75	0.64
152	The Disney character Donald is a	Het Disney-personage Donald is een	duck	eend	turkey	kalkoen	0.63	0.67
152	For our Christmas dinner. mother usually stuffed a	Voor ons kerstdiner vulde moeder gewoonlijk een	turkey	kalkoen	duck	eend	0.63	0.67
153	The American had a beautiful collection of birds of prey. but his favorite was his	De Amerikaan had een prachtige collectie roofvogels. maar zijn favoriet was zijn	eagle	arend	fly	vlieg	0.60	0.78
153	An insect that is attracted to shit is a	Een insect dat wordt aangetrokken door stront is een	fly	vlieg	eagle	arend	0.60	0.78
154	The dog buried a	De hond begroef een	bone	bot	heel	hak	0.64	0.73
154	To look taller she wore shoes with a	Om er langer uit te zien droeg ze schoenen met een	heel	hak	bone	bot	0.64	0.73
155	It is fashionable again to listen to music from a	Het is weer in de mode om muziek te luisteren van een	recordplayer	platenspeler	fan	fan	0.72	0.75
155	Messi signed the football for a	Messi tekende de voetbal voor een	fan	fan	recordplayer	platenspeler	0.72	0.75
156	The shepherd shaved a	De herder schoor een	sheep	schaap	farm	boerderij	0.58	0.66
156	They raised pigs on their	Ze fokten varkens op hun	farm	boerderij	sheep	schaap	0.58	0.66
157	Ana accidentally tripped and fell down the	Anna struikelde per ongeluk en viel van de	stairs	trap	fence	hek	0.77	0.68
157	To keep the dogs in the yard he put up a	Om de honden in de tuin te houden plaatste hij een	fence	hek	stairs	trap	0.77	0.68
158	One year after her death. Bill visited his mother’s	Een jaar na haar dood bezocht Bill zijn moeders	grave	graf	flower	bloem	0.66	0.71
158	There was a butterfly on a	Er zat een vlinder op een	flower	bloem	grave	graf	0.66	0.71
159	He always looked sharp with his suit and his	Hij zag er altijd netjes uit met zijn pak en zijn	tie	das	coat	jas	0.64	0.60
159	Let me take your hat and your	Laat me je hoed aannemen en je	coat	jas	tie	das	0.64	0.60
160	Dinner was not served in a bowl but on a	De maaltijd werd niet geserveerd in een kom maar op een	plate	bord	fork	vork	0.66	0.63
160	He ate the sausage with a knife and	Hij at de worst met mes en	fork	vork	plate	bord	0.66	0.63

161	The stable boy scooped up the hay with a	De stalknecht schepte het hooi op met een	fork	hooivork	pinecone	dennenappel	0.70	0.67
161	On a branch of the needle-leaved tree grew a	Op een tak van de naaldboom groeide een	pinecone	dennenappel	fork	hooivork	0.70	0.67
162	Her right foot was cold and she took a	Haar rechtervoet was koud en ze pakte een	sock	sok	toothbrush	tandenborstel	0.74	0.62
162	Don’t forget your pajamas and your	Vergeet niet je pyjama en je	toothbrush	tandenborstel	sock	sok	0.74	0.62
163	The animal that can grow back his lost tail is called a	Het dier dat zijn verloren staart kan laten terug groeien heet een	lizard	hagedis	crab	krab	0.68	0.66
163	Surimi is not real	Surimi is geen echte	crab	krab	lizard	hagedis	0.68	0.66
164	He kept his lawn nice and tidy with his	Hij hield zijn grasveld mooi en netjes met zijn	lawnmower	grasmaaier	gas	benzine	0.70	0.71
164	I would drive. but my car is low on	Ik zou rijden. maar mijn auto heeft nog maar weinig	gas	benzine	lawnmower	grasmaaier	0.70	0.71
165	Covered with a white sheet. he looked like a	Bedekt met het witte laken zag hij eruit als een	ghost	spook	wolf	wolf	0.72	0.73
165	The three little pigs were afraid of a	De drie kleine biggetjes waren bang voor een	wolf	wolf	ghost	spook	0.72	0.73
166	He didn't own a computer. so he wrote his books on a	Hij had geen computer dus hij schreef zijn boeken op een	typewriter	typmachine	letter	brief	0.63	0.63
166	He wrote his parents a	Hij schreef zijn ouders een	letter	brief	typewriter	typemachine	0.63	0.63
167	The jockey hit his horse's flank with a	De jockey sloeg zijn paard op de flank met een	whip	zweep	lightning	bliksem	0.74	0.73
167	She was afraid of the thunder and	Ze was bang voor de donder en	lightning	bliksem	whip	zweep	0.74	0.73
168	The draftsman sharpened his	De tekenaar sleep zijn	pencil	potlood	lips	lippen	0.79	0.73
168	He kissed his lucky coin with his	Hij kuste zijn geluksmunt met zijn	lips	lippen	pencil	potlood	0.79	0.73
169	I was attracted to him like a	Ik voelde me aangetrokken tot hem als een	magnet	magneet	switch	schakelaar	0.79	0.66
169	Mohamed wanted to turn on the light but he found no	Mohamed wilde het licht aandoen maar hij vond geen	switch	schakelaar	magnet	magneet	0.79	0.66
170	The athlete won a gold	De atleet won een gouden	medal	medaille	trophy	beker	0.54	0.78
170	The team that wins the most matches receives a	Het team dat de meeste wedstrijden wint ontvangt een	trophy	beker	medal	medaille	0.54	0.78
171	The hand of the captain was eaten by a crocodile and was now replaced with a	De hand van de kapitein was opgegeten door een krokodil en was nu vervangen door een	hook	haak	knot	knoop	0.66	0.78
171	He tied the rope to the pole with a complex	Hij bond het touw aan de paal met een ingewikkelde	knot	knoop	hook	haak	0.66	0.78
172	The train conductor blew a	De conducteur blies op een	whistle	fluitje	microphone	micro	0.79	0.80
172	The audience can't hear you if you don't speak into the	Het publiek kan je niet horen als je niet spreekt door de	microphone	micro	whistle	fluitje	0.79	0.80
173	The car had a flat	De auto had een platte	tire	band	wheel	wiel	0.60	0.75
173	She took her bicycle and saw that there was a spoke missing in the	Ze pakte haar fiets en zag dat er een spaak miste in het	wheel	wiel	tire	band	0.60	0.75
174	The car had to stop at a	De auto moest stoppen bij een	trafficlight	licht	moon	maan	0.74	0.65
174	In 1969 Neil Armstrong travelled to the	In 1969 reisde Neil Armstrong naar de	moon	maan	trafficlight	licht	0.74	0.65
175	You can catch malaria if you are bitten by a	Je kunt malaria krijgen als je gestoken wordt door een	mosquito	mug	spider	spin	0.63	0.73

175	In the middle of the large web sat a	Midden in het grote web zat een	spider	spin	mosquito	mug	0.63	0.73
176	The wooden plank for the floor was made shorter with a	De houten plank voor de vloer werd korter gemaakt met een	saw	zaag	nail	spijker	0.77	0.62
176	The carpenter secured the shelf with another	De timmerman zette de plank vast met nog een	nail	spijker	saw	zaag	0.77	0.62
177	He hung his shirt in the closet on a	Hij hing zijn hemd in de kast op een	hanger	kapstok	mailbox	brievenbus	0.77	0.77
177	He found a postcard from Portugal in his	Hij vond een ansichtkaart uit Portugal in zijn	mailbox	brievenbus	hanger	kapstok	0.77	0.77
178	He filled the bucket and closed the	Hij vulde de emmer en sloot de	tap	kraan	plug	stekker	0.68	0.77
178	We cannot put the lamp there. There is no outlet for the	We kunnen de lamp daar niet neerzetten. Er is geen stopcontact voor de	plug	stekker	tap	kraan	0.68	0.77
179	The well-known artist took a block of marble and carved a	De bekende kunstenaar nam een blok marmer en hakte een	statue	beeld	puzzle	puzzel	0.80	0.78
179	It was almost finished; there was the last piece of his	Het was bijna klaar; daar was het laatste stukje van zijn	puzzle	puzzel	statue	beeld	0.80	0.78
180	The pan fell on top of a	De pan viel bovenop een	pot	pot	hand	hand	0.71	0.76
180	He held the gun in his right	Hij hield het pistool in zijn rechter	hand	hand	pot	pot	0.71	0.76
181	He is as proud as a	Hij is zo trots als een	peacock	pauw	frog	kikker	0.71	0.73
181	Close by the pond she heard the croaking of a little green	Vlakbij de vijver hoorde ze het gekwaak van een kleine groene	frog	kikker	peacock	pauw	0.71	0.73

APPENDIX 3B

Table B1

Estimates, standard errors, t-values and p-values for the fixed and random effects of the final general linear mixed effect model for the prediction time frame in the target data set.

	Elog			
Fixed effects	β	se	t	p
Intercept	-3.478	0.148	-23.510	<.001
Language	-0.191	0.055	-3.487	<.001
Image type (target vs. Unrelated)	-1.502	0.123	-12.227	<.001
Target onset time	0.248	0.063	3.925	<.001
List2	0.599	0.154	3.896	<.001
List 3	0.401	0.155	2.595	0.012
List 4	0.276	0.154	1.789	0.080
List 5	0.341	0.154	2.219	0.031
List 6	0.464	0.178	2.601	0.012
List 7	0.736	0.160	4.597	<.001
List 8	0.112	0.160	0.699	0.488
Language:Image type	0.262	0.077	3.398	0.001
Image type: Target onset time	-0.385	0.087	-4.424	<.001
	Variance	SD		
Random effects				
Sentence				
(intercept)	1.563	1.250		
Image type	2.924	1.710		
Participant				
(intercept)	0.242	. 0.492		
Image type	0.209	0.456		

Table B2

Estimates, standard errors, t-values and p-values for the fixed and random effects of the final general linear mixed effect model for the prediction time frame in the competitor data set.

	Elog			
Fixed effects	β	se	*t*	*p*
Intercept	-4.172	0.154	-27.177	<.001
Language	-0.071	0.054	-1.313	0.189
Image type (Competitor vs. Unrelated)	-0.658	0.104	-6.349	<.001
Semantic distance	-0.186	0.054	-3.451	0.001
Target onset time	0.208	0.061	3.432	0.001
List2	0.629	0.179	3.515	0.001
List 3	0.500	0.180	2.784	.0.008
List 4	0.391	0.179	2.182	0.034
List 5	0.444	0.179	2.482	0.016
List 6	0.567	0.208	2.728	0.009
List 7	0.849	0.186	4.561	<.001
List 8	0.145	0.186	0.778	0.440
Language:Image type	0.144	0.076	1.895	0.058
Image type: Semantic distance	0.223	0.073	3.037	0.002
Language: Semantic distance	0.159	0.055	2.905	0.004
Image type: Target onset time	-0.291	0.082	-3.568	<.001
Language: Image type: Semantic distance	-0.191	0.077	-2.487	0.013
	Variance	SD		
Random effects				
Sentence				
(intercept)	1.378	1.174		
Image type	2.431	1.559		
Participant				
(intercept)	.0.158	0.397		
Image type	0.063	0.251		

APPENDIX 3C

Repeated testing (in each time bin) increases the likelihood of Type I errors. To show that the pattern of results remains the same with a Bonferroni corrected alpha value we plot the p-values of the most relevant effects in each time bin. Figure C1 shows the target data p-values of the interaction between language and image type in each bin.

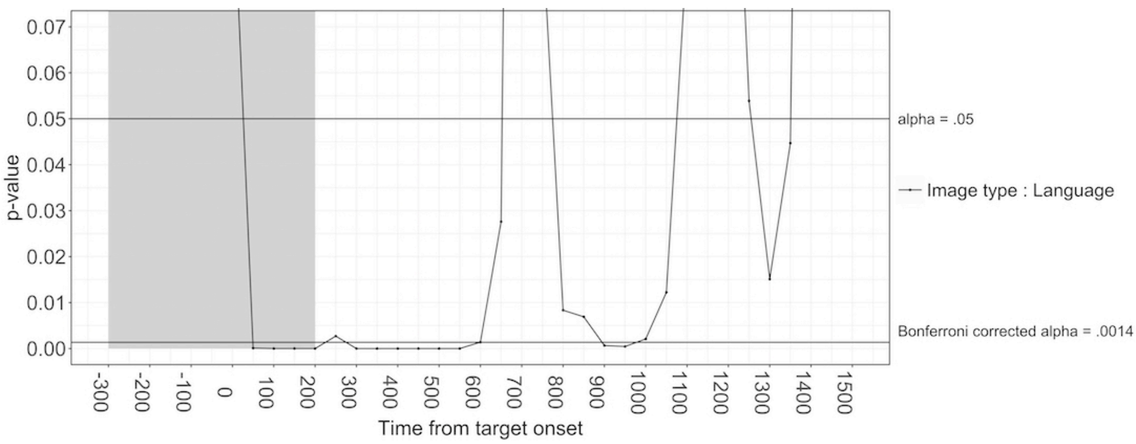


Figure C1. P-values of the language by image type interaction in each time bin (optimal model). Display onset was 500 ms before target onset. The area shaded grey is the prediction time frame. Horizontal lines indicate uncorrected alpha (0.05), and bonferroni corrected alpha (0.0014).

Figure C1 shows that the Image type by Language interaction is significant in the same time bins in the prediction time frame if we use Bonferroni corrected alpha (from 50 ms after target onset). The interaction remains significant until the time bin of 600-650 ms after target onset (except for 250-300 ms bin).

Figure C2 shows the competitor data p-values of the effects listed in the legend in each time bin.

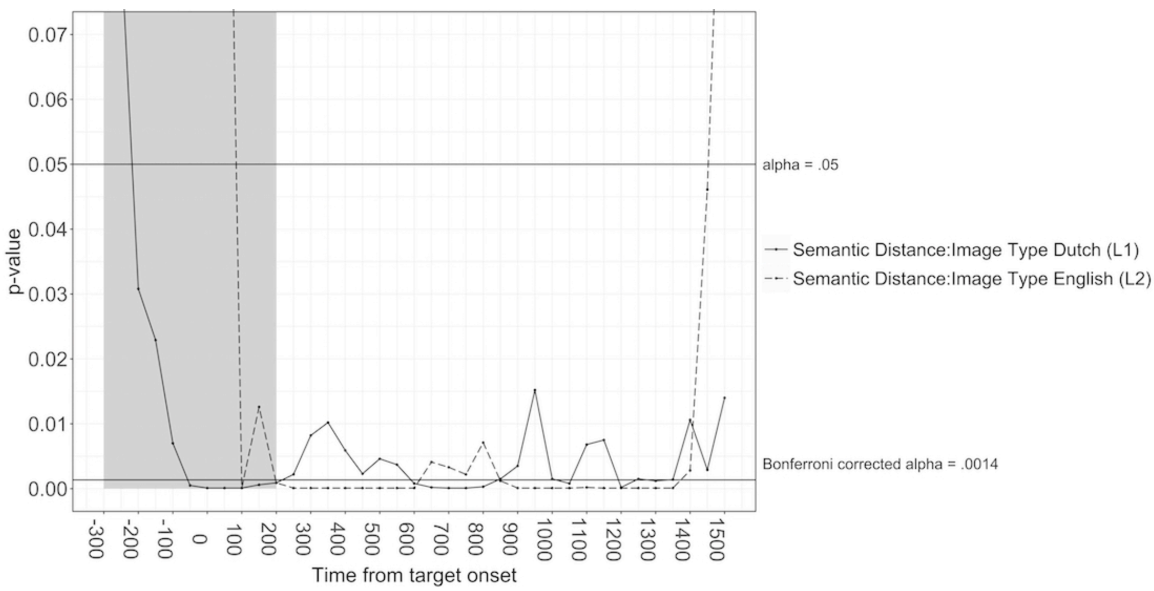
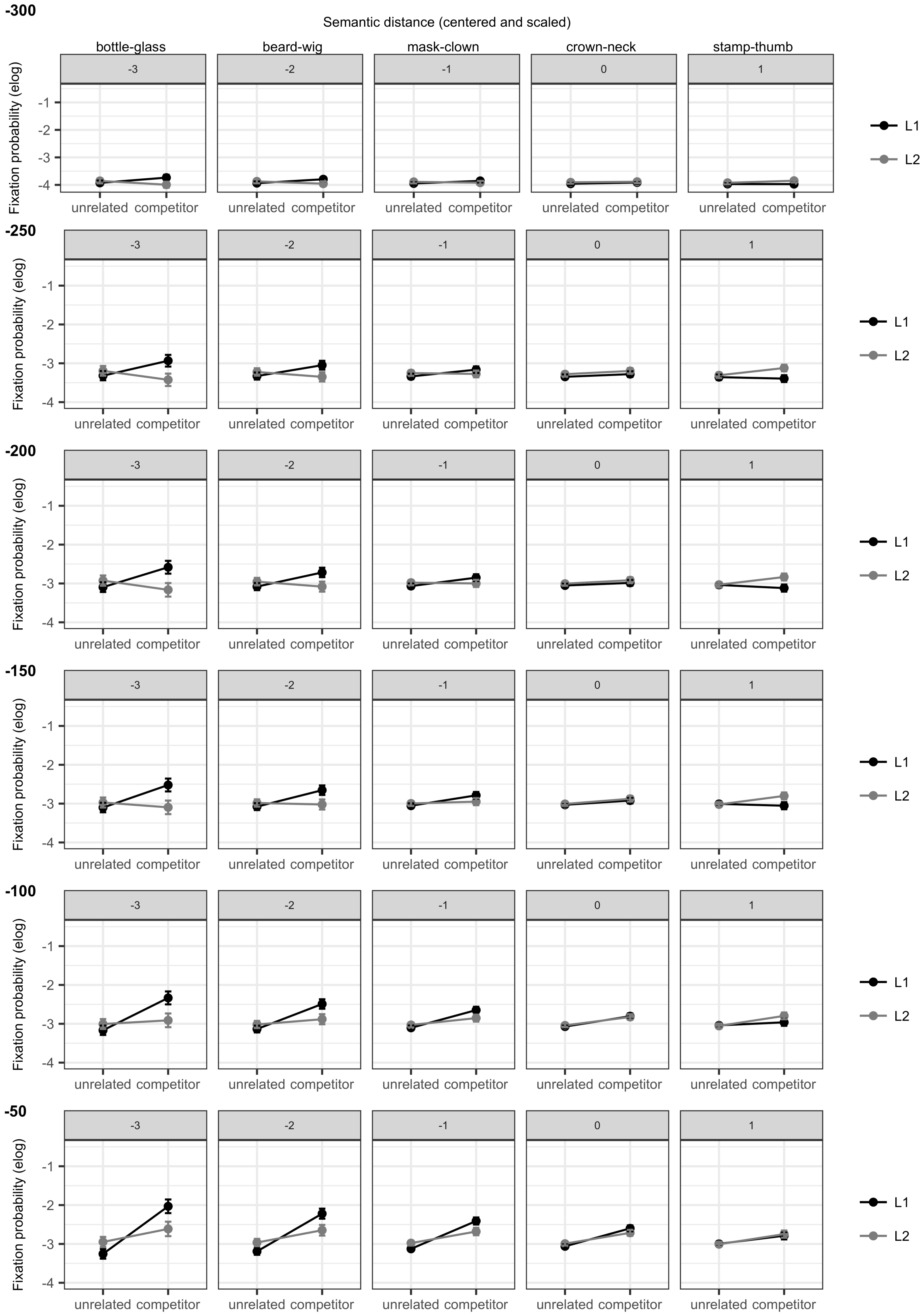
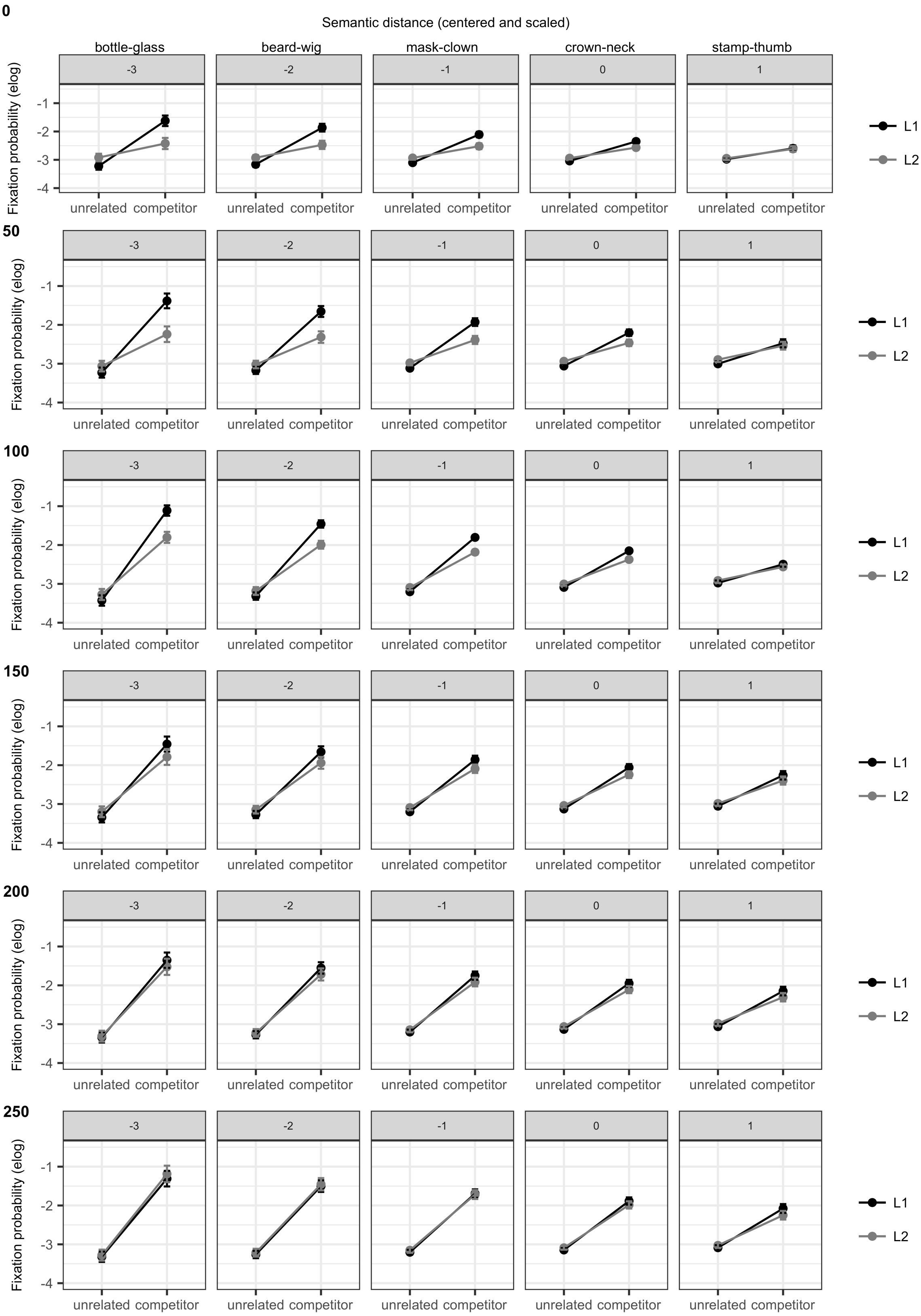


Figure C2. P-values of the effects in each time bin (optimal model). Display onset was 500 ms before target onset. The area shaded grey is the prediction time frame. Horizontal lines indicate uncorrected alpha (0.05) and Bonferroni corrected alpha (0.0014).

Figure C2 shows that the interaction between image type and semantic distance becomes significant 3 time bins later in Dutch (L1) if we use Bonferroni corrected alpha. However, in English (L2) there is still a delay of three time bins before the interaction becomes significant for the first time, and the interaction is consistently signicant from 200 ms after target onset (after the prediction time frame). Thus, the main pattern of results found with corrected alpha is the same as the pattern found with uncorrected alpha.

APPENDIX 3D


-300
-250
-200
-150
-100
-50
Semantic distance (centered and scaled)
bottle-glass
beard-wig
mask-clown
crown-neck
stamp-thumb
-3
-2
-1
0
1
Fixation probability (elog)
unrelated
competitor
L1
L2


0
50
100
150
200
250
Semantic distance (centered and scaled)
bottle-glass
beard-wig
mask-clown
crown-neck
stamp-thumb
-3
-2
-1
0
1
Fixation probability (elog)
-1
-2
-3
-4
unrelated competitor
L1
L2

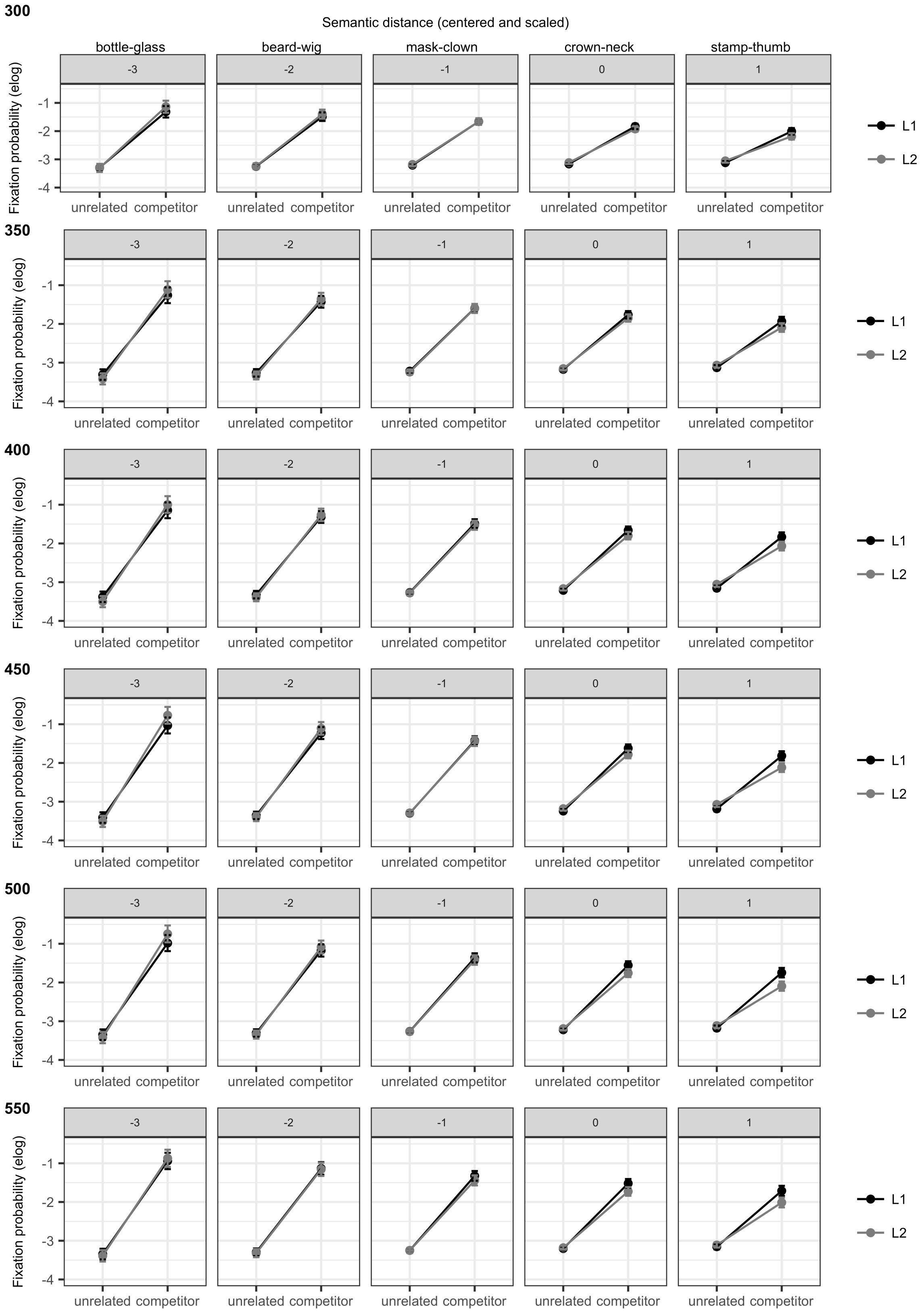


Figure D1. Three-way interaction between image type, language and semantic distance per time bin. Plot label in the left upper corner of each plot indicates time relative to target onset. The word pairs above each semantic distance facet are example competitor-word pairs for that semantic distance score.

APPENDICES CHAPTER 4

APPENDIX 4A

Table A1

Factors and interactions included in the full model for the Target trials and Competitor trials

Factors	**Two-way interactions**	**Three-way interactions**	**Four-way interactions**
Language (L1 Dutch vs. L2 English)	*Language * Image type*	*Language *Image type *Condition*	Language * Image type * Condition * Digit span
	Language Condition*	Language* Image type* Cloze probability	
	Language * Sentence syllable count	Language * Image type * English LexTALE score	
	Language * Cloze probability	Language*Image type* Digit span forward	
	Language* Digit span forward	Language *Image type * phonetic similarity	
	Language* Block	Language*Image type* Block	
	Language* Experimental image frequency	Language* Image type* Expeirmental image frequency	
	Language* Experimental image phonetic similarity	Language* Image type* Sentence syllable count	
	Language * English LexTALE score	Language* Condition* Digit span forward	
Image type (experimental vs. unrelated)	*Image type* Condition*	Image type * Condition * Digit span	
	Image type* Cloze probability		
	Image type* Digit span forward		
	Image type* Block		
	Image type* Picture repetition		
	Image type* Experimental image frequency		
	Image type* Experimental image phonetic similarity		

	Image type* Experimental image phoneme count		
	Image type* Sentence syllable count		
	Image type* English LexTALE score		
Condition	Condition* Digit span forward		
Sentence syllable count (upto target word)			
Cloze probability			
Block			
Picture repetition			
Digit span forward			
Presentation list			
English LexTALE score			
Experimental image frequency			
Experimental image phoneme count			
Experimental image phonetic similarity (Levenshtein distance between L1 and L2 translation equivalents)			
Additional terms competitor model			
Fixed factors	**Two-way interactions**	**Three-way interactions**	**Four-way interactions**
Semantic distance (between competitor and target, continuous variable)	*Language* Semantic distance*	*Language * Image type* Semantic distance*	*Language*Image type*Semantic Distance* Condition*
	*Image type * Semantic distance*	*Language* Semantic distance* Condition*	
	Condition Semantic distance*	*Image type* Semantic distance* Condition*	
Plausibility (plausibility rating of competitor word as sentence ending)	Language * Plausibility	Language * Image type* Plausibility	
	Image type * Plausibility		
Competitor grammaticality (as sentence ending)	Image type* Competitor grammaticality	Language* Image type* Competitor grammaticality	

	Language* Competitor grammaticality	

Note. Main experimental terms in italics. Secondary terms in normal font. Main experimental terms were never removed from the models. Condition refers to the factor Load (load vs. no load) in Experiment 1. In Experiment 2, condition refers to speed manipulation (manipulated vs. non-manipulated).

APPENDIX 4B

Table B1

Results final model target trials Experiment 1

Fixed effects	β	se	t	p
(Intercept)	-3.178	.099	-32.157	<.001
Language	-.394	.062	-6.358	.000
Image type	-1.484	.122	-12.144	<.001
Digit span forward	.021	.069	.301	.764
Experimental picture repetition	-.149	.024	-6.157	.000
Load	-.447	.063	-7.067	.000
Phonetic similarity	.057	.070	.818	.414
Language*Image type	.475	.087	5.430	.000
Image type*Experimental picture repetition	.132	.032	4.100	.000
Image type*Load	.460	.087	5.298	.000
Image type*Digit span forward	-.029	.075	-.390	.697
Image type* Phonetic similarity	-.058	.096	-.608	.543
Language*Load	.219	.087	2.509	.012
Language*Digit span forward	-.041	.031	-1.300	.194
Language*Phonetic similarity	.129	.044	2.943	.003
Digit span forward*Load	-.122	.046	-2.645	.008
Language*Image type*Phonetic similarity	-.185	.062	-2.986	.003
Image type*Digit span forward*Load	.227	.062	3.653	.000
Language*Image type*Load	-.273	.123	-2.222	.026

	Variance	*SD*
Item		
Intercept	1.409	1.187
Image type	2.642	1.626
Participant		
Intercept	.282	.531
Image type	.273	.522
Load	.015	.124
Experimental picture repetition	.004	.061

Table B2

Target trial significance levels in each time bin Experiment 1

	Time relative to target onset																
	Prediction time frame										**Post-prediction (target) time frame**						
Effects	**-300**	**-250**	**-200**	**-150**	**-100**	**-50**	**0**	**50**	**100**	**150**	**200**	**250**	**300**	**350**	**400**	**450**	**500**
(Intercept)																	
Language																	
Image type																	
Digit span forward																	
Experimental picture repetition																	
Load																	
Phonetic similarity																	
Language*Image type																	
Image type*Experimental picture repetition																	
Image type*Load																	
Image type*Digit span forward																	
Image type* Phonetic similarity																	
Language*Load																	
Language*Digit span forward																	
Language*Phonetic similarity																	
Digit span forward*Load																	
Language*Image type*Phonetic similarity																	

Image type*Digit span forward*Load	
Language*Image type*Load	

p>.1	.1>p>.05	.05>p>.01	.01>p>.001	p<.001

Table B3

Results final model competitor trials Experiment 1

Fixed effects	β	se	t	p
(Intercept)	-3.833	.090	- 42.718	<
Language	-.164	.062	-2.622	.009
Image type (competitor vs. unrelated)	-.505	.110	-4.593	.000
Block	-.036	.045	-.787	.431
Load	-.112	.061	-1.829	.067
Semantic distance	-.020	.055	-.367	.714
Language* Image type	.192	.087	2.219	.026
Image type* Block	-.080	.061	-1.305	.192
Image type* Load	.021	.086	.243	.808
Image type* Semantic distance	.060	.076	.783	.434
Language* Load	-.006	.086	-.070	.944
Language* Semantic distance	-.060	.063	-.952	.341
Load* Semantic distance	.028	.060	.460	.645
Language*Image type*Load	-.016	.122	-.132	.895
Language*Image type*Semantic distance	.061	.089	.680	.496
Language*Load*Semantic distance	.004	.087	.047	.962
Image type*Load*Semantic distance	-.036	.085	-.426	.670
Language*Image type*Load*Semantic distance	-.044	.122	-.360	.719
	Variance	SD		
Item				
intercept	1.253	1.119		

Image type	2.268	1.506
Participant		
Intercept	.183	.428
Experimental picture occurence	.009	.093
Image type	.077	.277

Table B4

Competitor trial significance levels in each time bin Experiment 1

Time relative to target onset

	Prediction time frame										**Post-prediction (target) time frame**						
Effects	**-300**	**-250**	**-200**	**-150**	**-100**	**-50**	**0**	**50**	**100**	**150**	**200**	**250**	**300**	**350**	**400**	**450**	**500**
(Intercept)																	
Language																	
Image type (competitor vs. unrelated)																	
Block																	
Load																	
Semantic distance																	
Language* Image type																	
Image type* Block																	
Image type* Load																	
Image type* Semantic distance																	
Language* Load																	
Language* Semantic distance																	
Load* Semantic distance																	
Language*Image type*Load																	
Language*Image type*Semantic distance																	
Language*Load*Semantic distance																	

Image type*Load*Semantic distance			
Language*Image type*Load*Semantic distance			

p>.1	**.1>p>.05**	**.05>p>.01**	**.01>p>.001**	**p<.001**

Table B5

Results final model target trials Experiment 2

Fixed effects	β	se	t	p
(Intercept)	-3.698	.116	-31.974	<.001
Language	.278	.115	2.414	.017
Image type	-1,130	.129	-8.748	<.001
Cloze probability	.141	.043	3.303	.001
Block	.380	.109	3.498	.001
Picture repetition	-.079	.032	-2.514	.012
Speed condition	.172	.061	2.817	.005
Digit span forward	.025	.054	.47	.640
Phonetic similarity	.127	.068	1.867	.063
Sentence syllable count	.172	.053	3.239	.001
Image type* Language	-.087	.089	-.972	.331
Image type* Cloze probability	-.170	.058	-2.925	.003
Image type* Block	-.224	.085	-2.635	.008
Image type* picture occurence	.093	.044	2.096	.036
Image type* speed condition	-.228	.086	-2.649	.008
Image type* phonetic similarity	-.154	.093	-1.661	.097
Image type* sentence syllable count	-.232	.072	-3.233	.001
Language *block	-.484	.168	-2.874	.005
Language * speed condition	-.216	.086	-2.517	.012
Language * phonetic similarity	-.138	.043	-3.198	.001
Language * image type * phonetic similarity	.161	.061	2.648	.008
Language*image type* speed condition	.236	.121	1.946	.052

	Variance	SD
Item	1.336	1.156
Image type	2.432	1.559
Participant	.314	.560
Image type	.320	.565
Block	.073	.270
Sentence syllable count	.002	.049
Digit span forward	.074	.271

Table B6

Target trial significance levels in each time bin Experiment 2

Time relative to target onset

	Prediction time frame										**Post-prediction (target) time frame**						
Effect	**-300**	**-250**	**-200**	**-150**	**-100**	**-50**	**0**	**50**	**100**	**150**	**200**	**250**	**300**	**350**	**400**	**450**	**500**
(Intercept)																	
Language																	
Image type																	
Cloze probability																	
Block																	
Picture repetition																	
Speed condition																	
Digit span forward																	
Phonetic similarity																	
Sentence syllable count																	
Image type* Language																	
Image type* Cloze probability																	
Image type* Block																	
Image type* picture occurence																	
Image type* speed condition																	
Image type* phonetic similarity																	
Image type* sentence syllable count																	
Language *block																	
Language * speed condition																	
Language * phonetic similarity																	

Language * image type * phonetic similarity	
Language*image type* speed condition	

p>.1	.1>p>.05	.05>p>.01	.01>p>.001	p<.001

Table B7

Results final model competitor trials Experiment 2

Fixed effects	β	se	t	p
(Intercept)	-4.125	.116	-35.577	<.001
Language	.023	.128	.181	.857
Image type	-.538	.114	-4.731	.000
Block	.381	.120	3.185	.002
Experimental image occurence	-.112	.032	-3.450	.001
Speed condition	-.065	.060	-1.084	.279
Experimental image phoneme count	.058	.040	1.477	.140
Experimental image frequency	.022	.019	1.164	.244
Sentence syllable count	.156	.053	2.960	.003
Competitor plausibility	.140	.052	2.687	.007
Semantic distance	-.045	.055	-.820	.412
Image type*Language	.197	.089	2.202	.028
Image type*Block	-.183	.083	-2.195	.028
Image type*Experimental image occurence	.107	.044	2.435	.015
Image type* speed condition	.092	.085	1.093	.275
Image type* Experimental image phoneme count	-.068	.053	-1.283	.200
Image type* Sentence syllable count	-.220	.071	-3.090	.002

Image type*Competitor plausibility	-.212	.070	-3.013	.003
Image type* Semantic distance	.101	.076	1.326	.185
Language*Block	-.456	.203	-2.248	.027
Language*Speed condition	.184	.085	2.177	.029
Language*Semantic distance	.018	.062	.293	.770
Speed condition*Semantic distance	.027	.059	.458	.647
Language*Image type*Semantic distance	-.119	.087	-1.363	.173
Language*Image type*Speed condition	-.273	.120	-2.281	.023
Language*Speed condition*Semantic distance	.049	.085	.579	.563
Image type*Speed condition*Semantic distance	-.094	.083	-1.130	.259
Language*Image type*Speed condition*Semantic distance	.038	.120	.316	.752
Item	1.332	1.154		
Image type	2.386	1.545		
Participant	.295	.543		
Image type	.072	.269		
Experimental picture occurence	.005	.069		
Block	.034	.184		

Table B8

Competitor trial significance levels in each time bin Experiment 2

Time relative to target onset

	Prediction time frame										**Post-prediction (target) time frame**						
Effects	**-300**	**-250**	**-200**	**-150**	**-100**	**-50**	**0**	**50**	**100**	**150**	**200**	**250**	**300**	**350**	**400**	**450**	**500**
(Intercept)																	
Language																	
Image type																	
Block																	
Experimental image occurence																	
Speed condition																	
Experimental image phoneme count																	
Experimental image frequency																	
Sentence syllable count																	
Competitor plausibility																	
Semantic distance																	
Image type*Language																	
Image type*Block																	
Image type*Experimental image occurence																	
Image type* speed condition																	

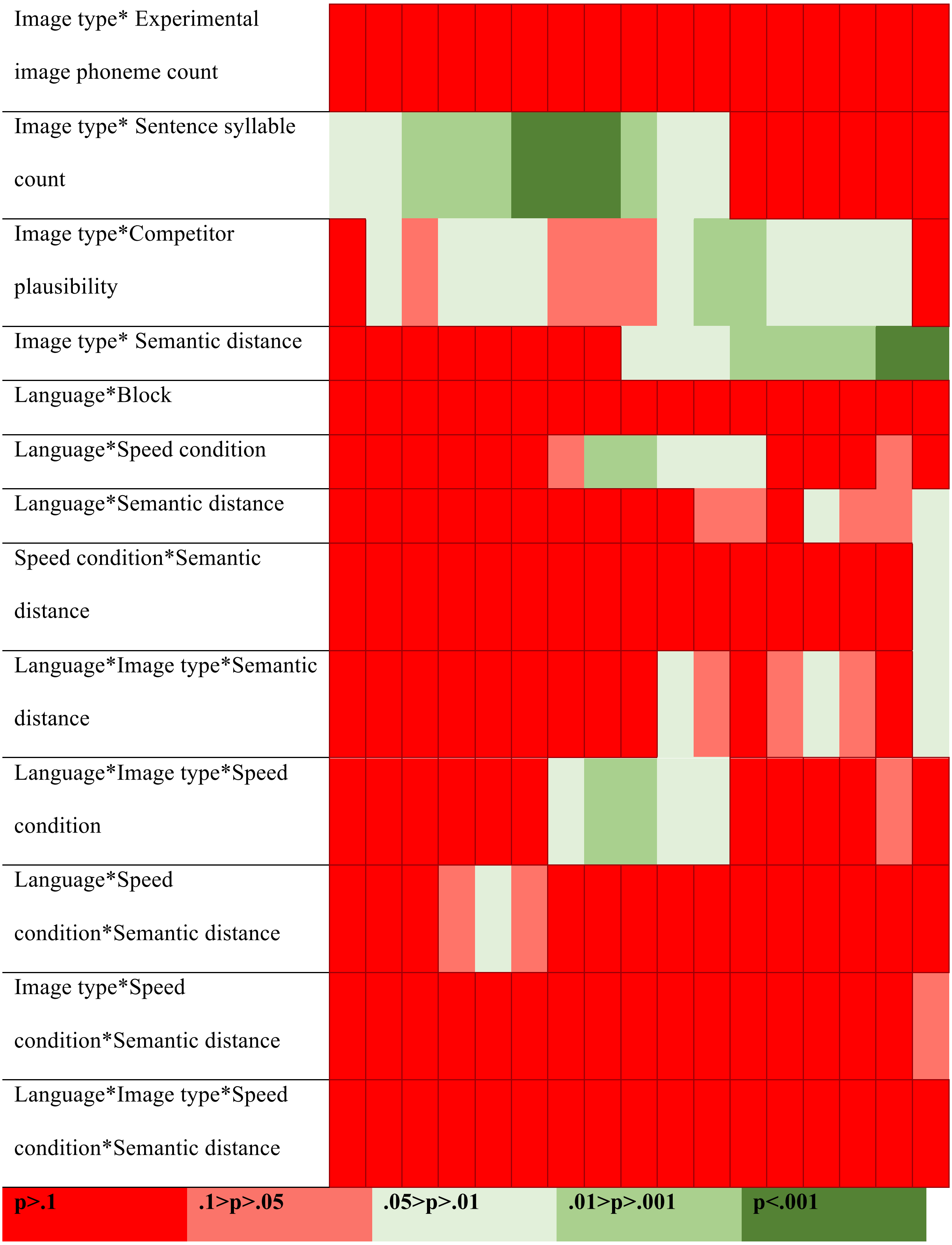


Image type* Experimental image phoneme count
Image type* Sentence syllable count
Image type*Competitor plausibility
Image type* Semantic distance
Language*Block
Language*Speed condition
Language*Semantic distance
Speed condition*Semantic distance
Language*Image type*Semantic distance
Language*Image type*Speed condition
Language*Speed condition*Semantic distance
Image type*Speed condition*Semantic distance
Language*Image type*Speed condition*Semantic distance
p>.1
.1>p>.05
.05>p>.01
.01>p>.001
p<.001

Appendices Chapter 5

Appendix 5A

Stimuli sentences and target words. When no target word is specified in the second or third column, the sentence in the first column is a filler sentence.

Sentence	**Target Word /æ/**	**Target Word /b/**
The Russian mob of New York was glad the police did not arrest them.	glad	mob
They prescribe a type of medicine that decreases gas in your bowels.	gas	prescribe
We rob all people with a hammer, said the criminal.	hammer	rob
The rich snob often paints a portrait of a landscape outside.	landscape	snob
The man was sitting on a stub while thinking about his future.	man	stub
While being in the pub on Mainstreet, he tends to slap people.	slap	pub
The woman decided to show a boob on the tram in the city center.	tram	boob
This band tours around the globe every two years.	band	globe
A friend of mine broke his rib on his left side due to a bat on the baseball field.	bat	rib
A tube of sand was used during the experiment.	sand	tube
Suzanne's job in the music industry was to rap on stage.	rap	job
Much of the fat was reduced with a probe inserted into the tissue by a doctor.	fat	probe
The panther lay on a stone in the form of a cube in the jungle.	panther	cube

Sentence	**Target Word /æ/**	**Target Word /b/**
He told me to rub a lamp to see a genie.	lamp	rub
I always enjoyed it when I had to dub a movie.	had	dub
Either choose a robe or a mantle, but not both.	mantle	robe
He felt a throb in his head due to the scam of the criminal.	scam	throb
They plan to bribe all the supervisors of the company.	plan	bribe
Melissa keeps one hand in the hot tub only because she likes the warmth.	hand	tub
All she did was sob in the shadow of the tree.	shadow	sob
The club in Denver purchased a car ramp for the parking lot.	ramp	club
The knob on the door in the old building was flat like a leaf.	flat	knob
The cub of the cat was too tired to play.	cat	cub
A web of a spider is its best trap to hunt its prey.	trap	web
Sergio always forgets to scrub around the gap in the floor.	gap	scrub
They organised a sports match with the tribe of Indians in the morning.	match	tribe
The hat of the old woman was covered with a blob of bird poop.	hat	blob
Bob often showed her a map of the subway.	map	bob
The babe in the cradle loves to play with the small pan in the kitchen.	pan	babe
I scan the crib in order to find little Lisa's favorite toy.	scan	crib
Her plan was to expose a boob on stage.	plan	boob
They needed a hammer to open the knob on the door.	hammer	knob
There was a shadow of the king's robe on the road.	shadow	robe

Sentence	Target Word /æ/	Target Word /b/
It is very difficult to dub a rap in a movie.	rap	dub
The teenager saw a total babe entering the tram to the center.	tram	babe
He knew the man loved to go to the club in London to perform.	man	club
There was a spider web on the old fur mantle in my mother's closet.	mantle	web
The guy lost his job of course, since he refused to remove his hat when serving customers.	hat	job
He knew it was a trap when Bob ordered him to lock the door.	trap	Bob
He bruised his rib in June because he did not notice a gap in the street.	gap	rib
He put a cube of butter into the pan to melt.	pan	cube
The big bat from the cave bit the poor lion cub only out of fear.	bat	cub
I think I had a stub of a pencil in my drawer somewhere.	had	stub
The skateboarder preferred the tube over the ramp since it was much more exciting.	ramp	tube
The cat enjoys it when you rub its stomach.	cat	rub
The mob in Sicily is involved in the theft of gas from cars.	gas	mob
He stepped out of the tub in order to observe the landscape through a window.	landscape	tub
I'm glad because I will never need to talk to this snob again.	glad	snob

Sentence	Target Word /æ/	Target Word /b/
The tribe occupying the strip of desert used sand to clean their pots.	sand	tribe
The officer showed the suspect the map in order to probe into what really occurred.	map	probe
The artist used his hand to remove a blob of paint.	hand	blob
The doctor needed to prescribe a number of drugs to the fat patient because he was diabetic.	fat	prescribe
It was an awful scam to try and sell the pub on the block which would be demolished.	scam	pub
Please turn on the lamp so I will be able to find my country on the globe in the corner.	lamp	globe
Since he expected his brother to rob a neighbor's flat he called the police.	flat	rob
The drummer of the band was told to scrub all of the dirt off of the stage.	band	scrub
The zoo keeper couldn't hear the panther's heart throb in his chest.	panther	throb
Sometimes he told her he would slap her if she would sob in public.	slap	sob
He made a fire using a match next to the crib in the nursery.	match	crib
She urged me to send a scan of the article on how to bribe a teacher.	scan	bribe
A large gas explosion occurred in the shop.	gas	
Much whisky was drunk on the party instead of beer.		

Sentence	Target Word /æ/	Target Word /b/
Her cleavage revealed a perfect boob in a pretty red bra.		boob
She took her boob out of her shirt in order to feed her baby.		boob
I don't want you to probe into my business.		probe
Shell wants to start searching for gas in the North Pole.	gas	
There was a huge gap between his teeth.	gap	
To bridge a gap, the directors paid the employees more.	gap	
He spilled some wine on her dress.		
Gently insert the probe into the mouth when the patient is asleep.		probe
The police arrested important members of the Chinese mob in their homes.		mob
He waited desperately for the lord's sign because he did not know what to do.		
My mother uses the large pan to cook the meat.	pan	
The pan caused a fire in the kitchen.	pan	
I was glad the problem could be solved.	glad	
The mob in New York is increasing its power in some neighborhoods.		mob
The web of the tiny spider reached all the way to the other side of the porch.		web
Don't get caught up in his web of lies again.		web
All tennis balls were collected by the assistant.		
You should be glad he did not sue you.	glad	
Desert sand is able to get inside your watch.	sand	

Sentence	Target Word /æ/	Target Word /b/
In winter, the children go out to play in the snow.		
The red wine stain may disappear if you rub a bit of salt on it.		rub
It would be great if you could rub a bit of sun block on my shoulders.		rub
I'd love to own a house with a tub in the bedroom.		tub
A lot of sand is used for the new garden.	sand	
Thor is armed with a large hammer according to myth.	hammer	
A yellow hammer is a kind of bird.	hammer	
Ben is too young to be a lawyer.		
My uncle built a tub in his own yard.		tub
The girl tried to bribe an officer in the parking lot.		bribe
There was a big explosion in Syria because of terrorists.		
Bobby's right hand was scarred by the fire.	hand	
Would you give me a hand with this ceiling?	hand	
Nobody wears a hat these days.	hat	
It seemed like she wanted to bribe a lawyer but I'm not sure.		bribe
The patient's tongue was so swollen he needed to breathe through a tube in his throat.		tube
There was a tube in there connecting the vessel to another one.		tube
He kicked in the door with his heel.		
I take my hat off for this accomplishment.	hat	
The bear walked right into the trap of the hunter.	trap	
The book which stood on the shelf fell on the floor.		

Sentence	Target Word /æ/	Target Word /b/
The tribe of Indians dispersed in the woods to confuse the explorers.		tribe
The spiral shaped scar on his shoulder meant he was part of the tribe of hunters.		tribe
When I was a child there was a globe in my room with a light in it.		globe
This useless trap did not kill the prey.	trap	
Julia found a man on the street who was shot.	man	
The common man does not know much about neurobiology.	man	
Emma is talking about the tigers she saw today on her trip to the zoo.		
Let's spin the globe in order to find a nice location for our spring trip.		globe
Tomorrow in the spa we could use sea salt to scrub our skin.		scrub
People whisper when they do not want to be heard.		
The criminal continued his scam on the street.	scam	
A good scam deprives you of all your accessories.	scam	
The mantle of the king was far too short.	mantle	
The maid really needs to scrub all the restrooms before the guests arrive.		scrub
She wore a gorgeous robe accompanied by the perfect pumps.		robe
The bishop couldn't find his robe anywhere this morning.		robe
It was too hot to sit outside to drink coffee.		

Sentence	Target Word /æ/	Target Word /b/
We covered the wounded soldier in a mantle of silk.	mantle	
Our kitten resembled a panther when she hunted.	panther	
Her colleague told her about their new boss.		
We should encourage them not to dub all French movies in order to boost learning.		dub
You may know her voice because she is often paid to dub a movie.		dub
You've been behaving like a snob all week.		snob
A panther is hard to see in the dark.	panther	
The biker used the ramp during the race.	ramp	
Your ramp caught fire since it is made of wood.	ramp	
The computer broke down because of a virus.		
The waiter serving us yesterday was a snob anyway.		snob
The model would like the surgeon to remove a rib in order to look slimmer.		rib
Everyone thought the white elegant outfit of the bride was beautiful.		
She used to slap her in the face.	slap	
A hard slap is said to help you focus.	slap	
Blake needed a CT-scan to find the tumor.	scan	
A rib eye steak is what I love most in the world.		rib
She urged the doctor to prescribe a pill from a different company.		prescribe

Sentence	Target Word /æ/	Target Word /b/
This doctor does not prescribe any medicine for a cold with good reason.		prescribe
Everyone listens attentively to the guide talking about the old church.		
You will need a scan of this document.	scan	
All of the pirates sought the treasure map of the island.	map	
My father always wants to be the best in chess.		
I will be fired next week but I didn't really like my job anyway.		job
These days it is very difficult to find a job in my field.		job
They heard a sudden throb a second before the motor died.		throb
Only a map will show us the way out of this maze.	map	
When I was young I had a teddy bear called Charly.	had	
Did you say you had a house with a swimming pool?	had	
Her father loves to take his luxurious car for a spin.		
This morning the wound started to throb a little.		throb
The leopard left his cub alone to go on a hunt.		cub
He went to the shop to buy a new book.		
My brother joined a band in order to become popular.	band	
The lead singer in a band mostly determines its success.	band	
Those two always try to match their outfits.	match	
When you see a bear cub alone you need to be cautious because the mother will not be far.		cub

Sentence	**Target Word /æ/**	**Target Word /b/**
To get in through the door you need to turn the knob on the other side.		knob
I'm not sure how to open it, I don't find the knob on this window.		knob
Marc goes to the therapist living in a nearby village.		
I could never match her chess skills.	match	
The tram in the Hague makes me nauseous.	tram	
Uncle Jerry needs a dentist because his tooth hurts.		
Today either Bob or Marc will win a bike in the tournament.		bob
You did not mention Bob all of a sudden leaving his wife for another.		bob
I'll be out partying in the club on Times Square tonight.		club
A Belgian tram does not show its current location.	tram	
Suzy got fat because she ate too much junk food.	fat	
My neighbors' fat dog was regularly overfed.	fat	
He saw some money lying on the floor in front of him.		
Would you like to join the club of supporters?		club
I never once witnessed someone rob a store.		rob
The musical on Broadway was amazing.		
The lamp in the changing room was broken.	lamp	
Jacob's chamber was lit only by the lamp on his desk.	lamp	
Turkish people make flat bread in a great oven.	flat	
Please don't tell me you gave him permission to rob all of them.		rob

Sentence	Target Word /æ/	Target Word /b/
Where the little crib of the girl used to be, there was now a desk.		crib
The rock star owns a crib in Florida the size of Disney World.		crib
The concerned uncle comforted the toddler on his first day of school.		
People thought the world was flat in the middle ages.	flat	
A scary bat rested on the ceiling of the cave.	bat	
Her mother likes her new scarf very much.		
When his work is finished he goes to the pub in a village nearby.		pub
I bought an old pub in need of remodeling.		pub
Put a cube of ice on the wound to reduce the pain.		cube
Billy could never hold the bat the right way.	bat	
A lot of people rap, but only few possess skill.	rap	
His rap music united two competing neighborhoods.	rap	
Her sister is a successful model working in New York.		
Today in school we learned how to draw a cube in 3D.		cube
Where the soldier's arm used to be there was only a stub of about 3 inches long now.		stub
He interviewed the victim of the assault.		
I plan to finish my thesis next year.	plan	
If everything goes according to plan, we should win the cup.	plan	
Germany possesses a beautiful landscape, especially in the west.	landscape	

Sentence	Target Word /æ/	Target Word /b/
It is impolite to toss your cigarette stub on the ground.		stub
I saw the boy sob a long time when his mother said goodbye to him on his first day of school.		sob
She did not want to show her tears but she could not help but sob all evening.		sob
He tossed his broom on the floor because he was on strike.		
The artist got inspired by the landscape of Spain.	landscape	
Their annoying cat always walks in our garden.	cat	
I need some tissues to clean this mess.		
There was always a hot babe in the company of the movie star.		babe
He hoped to hold the babe in his arms for the first time before he went to bed.		babe
He found a mysterious blob of jelly in the dirty old fridge.		blob
Their cat loved to hunt mice.	cat	
He loves to lurk in the shadow of the school.	shadow	
My cousin always tries to catch her shadow on the street.	shadow	
The binder contained information on the budget.		
The blue blob over there is the lake she was talking about.		blob

APPENDIX 5B

Table 2

Raw values of F1 and F2 in Hz, vowel duration and closure duration in ms divided by speaker, social context, and block. The standard deviations are presented in parentheses.

Condition	F1 /æ/	F2 /æ/	Vowel duration /æ/	Vowel duration /b/	Closure duration /b/
Confederate absent, baseline	868 (113)	1873 (241)	124.58 (35.3)	131 (44)	60 (12)
Confederate absent, post-exposure	844 (108)	1821 (221)	123.67(38.36)	131 (48)	60 (16)
Confederate present, baseline	886 (86)	1820 (209)	125.66 (34.17)	134 (41)	63 (12)
Confederate present, post-exposure	858 (83)	1802 (197)	121.00 (37.14)	131 (46)	62 (14)
Participant absent, baseline	751 (79)	1833 (162)	104.16 (28.93)	101 (40)	90 (27)
Participant absent, post-exposure	760 (79)	1837 (160)	101.84 (27.48)	99 (37)	80 (23)
Participant present, baseline	749 (169)	1848 (221)	107.67 (32.54)	104 (38)	87 (30)
Participant present, post-exposure	759 (112)	1826 (180)	106.34 (34.53)	104 (38)	78 (24)

APPENDIX 5C

Table C1

Estimates, standard errors, t-values and p-values for the fixed and random effects of the final general linear mixed effect model for the dependent measures of /b/, participant data set.

	Closure duration				Preceding vowel duration			
Fixed effects	β	se	*t*	*p*	β	se	t	p
Intercept	.297	.018	16.140	<.001	.363	.021	17.396	<.001
block	-.019	.005	-3.585	<.001	.020	.004	4.556	.000
social context	-.018	.014	-1.331	.192	.005	.014	.372	.712
list 2	.028	.018	1.525	.138	-.032	.019	-1.656	.106
list 3	.020	.019	1.106	.277	-.039	.019	-2.077	.046
list 4	-.004	.019	-.200	.843	-.017	.019	-.873	.388
block*social context	.004	.007	.566	.571	-.002	.006	-.354	.723
	Variance	SD			Variance	SD		
Random effects								
sentence								
(intercept)	<.001	<.001			<.001	.005		
participant								
(intercept)	.001	.035			.001	.036		
word								
(intercept)	.004	.060			.006	.080		
block	/	/						
list 2	/	/			.001	.029		
list 3	/	/			<.001	.019		
list 4	/	/			.001	.026		

Table C2

Estimates, standard errors, t-values and p-values for the fixed and random effects of the final general linear mixed effect model for the dependent measures of /b/, confederate data set.

	Closure duration				Preceding vowel duration			
Fixed effects	β	se	*t*	*p*	β	se	t	p
Intercept	.218	.009	23.025	<.001	.468	.018	26.416	<.001
block	-.009	.008	-1.176	.240	-.008	.011	-.759	.449
social context	.001	.006	.138	.890	-.007	.008	-.881	.379
session	<.001	<.001	-.435	.664	-.001	<.001	-2.947	.003
list 2	-.003	.004	-.699	.485	.004	.007	.553	.581
list 3	-.007	.003	-2.565	.010	-.008	.004	-2.175	.030
list 4	-.011	.004	-2.771	.006	.004	.007	.502	.617
block*social context	.006	.008	.720	.472	-.001	.011	-.077	.938
block*session	<.001	<.001	.903	.367	<.001	<.001	1.255	.210
social context*session	<.001	<.001	1.719	.086	<.001	<.001	.618	.537
block*social context*session	<.001	<.001	-.764	.445	<.001	.001	-.843	.399
Random effects	Variance	SD			Variance	SD		
sentence								
(intercept)	<.001	.022			.001	.039		
social context	<.001	.007			<.001	.012		
word								
(intercept)	.002	.041			.007	.085		
social context	<.001	.009			<.001	.012		

Table C3

Estimates, standard errors, t-values and p-values for the fixed and random effects of the final general linear mixed effect model for the dependent measures of /æ/, participant data set.

	F1				F2				duration			
Fixed effects	β	se	*t*	*p*	β	se	t	p	β	se	t	p
Intercept	1.103	.018	62.632	<.001	.995	.009	111.811	<.001	.361	.022	16.603	<.001
block	.013	.009	1.361	.183	-.002	.002	-.674	.501	.012	.008	1.615	.114
social context	-.009	.016	-.547	.589	.002	.007	.279	.782	-.013	.016	-.780	.442
list 2	.006	.021	.280	.781	.015	.010	1.467	.152	-.035	.019	-1.811	.079
list 3	-.020	.021	-.962	.343	.002	.010	.206	.838	-.029	.019	-1.534	.135
list 4	-.018	.021	-.850	.402	.016	.010	1.592	.121	-.028	.019	-1.470	.151
block*social context	.011	.013	.839	.407	-.001	.003	-.438	.661	<.001	.010	-.026	.979
	Variance	SD			Variance	SD			Variance	SD		
Random effects												
sentence												
(intercept)	<.001	<.001			<.001	.005			<.001	.015		
block	/	/							/	/		
participant												
(intercept)	.002	.043			<.001	.020			.002	.045		
block	.001	.029							<.001	.022		
word												
(intercept)	.001	.023			<.001	.021			.006	.077		
block	<.001	.013							<.001	.004		
Social context	<.001	.013										
list 2	/	/			/	/			/	/		
list 3	/	/			/	/			/	/		
list 4	/	/			/	/			/	/		

Table C4

Estimates, standard errors, t-values and p-values for the fixed and random effects of the final general linear mixed effect model for the dependent measures of /æ/, confederate data set.

	F1				F2				duration			
Fixed effects	β	se	*t*	*p*	β	se	t	p	β	se	t	p
Intercept	1.124	.019	6.080	<.001	1.029	.011	89.764	<.001	.424	.021	19.789	<.001
block	.009	.008	1.095	.275	-.016	.005	-2.949	.004	.003	.006	.434	.665
social context	.010	.008	1.296	.197	-.023	.004	-6.145	<.001	-.027	.011	-2.481	.017
session	.001	<.001	2.868	.004	-.001	<.001	-5.616	<.001	/	/	/	/
list 2	.011	.005	2.222	.032	-.002	.003	-.488	.627	-.018	.015	-1.169	.249
list 3	-.002	.003	-.461	.647	.001	.002	.365	.715	-.005	.014	-.328	.745
list 4	.018	.005	3.530	.001	-.002	.003	-.678	.499	-.028	.015	-1.853	.071
block*social context	-.008	.009	-.890	.374	.008	.005	1.517	.130	.005	.006	.900	.369
block*session	-.001	<.001	-3.318	.001	<.001	<.001	.932	.352	/	/	/	/
social context*session	-.001	<.001	-4.122	.000	.001	<.001	6.491	<.001	/	/	/	/
block*social context*session	.001	<.001	2.676	.008	<.001	<.001	-.357	.721	/	/	/	/
	Variance	SD			Variance	SD			Variance	SD		
Random effects												
sentence												
(intercept)	.001	.024			<.001	.019	<.001	.019	.001	.026		
social context	<.001	.010			<.001	.008	<.001	.008	.000	.006		
participant												
(intercept)	/	/			/	/			.001	.027		
word												
(intercept)	.009	.097			.003	.059			.009	.097		
block	<.001	.015			<.001	.006			/	/		
social context	.001	.027			<.001	.006			<.001	.014		
list 2	<.001	.014										
list 3	<.001	.009										
list 4	<.001	.015										

DATA STORAGE FACT SHEET CHAPTER 2

% Data Storage Fact Sheet

% Chapter 2

% Author: Aster Dijkgraaf

% Date: 28 June 2018

1. Contact details

===

1a. Main researcher

- name: Aster Dijkgraaf
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: aster.dijkgraaf@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Wouter Duyck
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: wouter.duyck@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies

===

* Reference of the publication in which the datasets are reported:

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. Bilingualism: Language and Cognition, 20(05), 917,Äì930. https://doi.org/10.1017/S1366728916000547

* Which datasets in that publication does this sheet apply to?:

All data sets reported in the publication

3. Information about the files that have been stored

===

3a. Raw data

----

* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

If NO, please justify:

* On which platform are the raw data stored?
  - [X] researcher PC
  - [ ] research group file server
  - [X] other (specify): external harddrive

* Who has direct access to the raw data (i.e., without intervention of another person)?
  - [X] main researcher
  - [ ] responsible ZAP
  - [ ] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files

----

* Which other files have been stored?
  - [ ] file(s) describing the transition from raw data to reported results. Specify:
  - [X] file(s) containing processed data. Specify: csv files used for the analyses
  - [X] file(s) containing analyses. Specify: R script
  - [ ] files(s) containing information about informed consent
  - [ ] a file specifying legal and ethical provisions
  - [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...
  - [ ] other files. Specify: ...

* On which platform are these other files stored?
  - [X] individual PC
  - [ ] research group file server
  - [X] other (specify): external harddrive

* Who has direct access to these other files (i.e., without intervention of another person)?
  - [X] main researcher
  - [ ] responsible ZAP
  - [ ] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction

===

* Have the results been reproduced independently?: [ ] YES / [X] NO

* If yes, by whom (add if multiple):
  - name:
  - address:
  - affiliation:
  - e-mail:

v0.2

DATA STORAGE FACT SHEET CHAPTER 3

% Data Storage Fact Sheet

% Chapter 3
% Author: Aster Dijkgraaf
% Date: 28 June 2018

1. Contact details
==

1a. Main researcher

- name: Aster Dijkgraaf
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: aster.dijkgraaf@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Wouter Duyck
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: wouter.duyck@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
==
* Reference of the publication in which the datasets are reported:

Dijkgraaf, Hartsuiker, and Duyck (Submitted). Prediction and Integration of Semantics during L2 and L1 Listening.

* Which datasets in that publication does this sheet apply to?:

All data sets reported in the publication

3. Information about the files that have been stored
===

3a. Raw data
----

* Have the raw data been stored by the main researcher? [X] YES / [ ] NO
If NO, please justify:

* On which platform are the raw data stored?
  - [X] researcher PC
  - [ ] research group file server
  - [X] other (specify): external harddrive

* Who has direct access to the raw data (i.e., without intervention of another person)?
  - [X] main researcher
  - [ ] responsible ZAP
  - [ ] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files
----

* Which other files have been stored?
  - [ ] file(s) describing the transition from raw data to reported results. Specify:
  - [X] file(s) containing processed data. Specify: csv files used for the analyses
  - [X] file(s) containing analyses. Specify: R script
  - [ ] files(s) containing information about informed consent
  - [ ] a file specifying legal and ethical provisions
  - [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...
  - [ ] other files. Specify: ...

* On which platform are these other files stored?
  - [X] individual PC
  - [ ] research group file server
  - [X] other (specify): external harddrive, and on open science framework: osf.io/wy9tm

```
* Who has direct access to these other files (i.e., without intervention of another person)?
  - [X] main researcher
  - [ ] responsible ZAP
  - [ ] all members of the research group
  - [ ] all members of UGent
  - [X] other (specify): anyone


4. Reproduction
============================================================
* Have the results been reproduced independently?: [ ] YES / [X] NO

* If yes, by whom (add if multiple):
   - name:
   - address:
   - affiliation:
   - e-mail:


v0.2
```

DATA STORAGE FACT SHEET CHAPTER 4

```
% Data Storage Fact Sheet

% Chapter 4
% Author: Aster Dijkgraaf
% Date: 28 June 2018


1. Contact details
============================================================

1a. Main researcher
------------------------------------------------------------
- name: Aster Dijkgraaf
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: aster.dijkgraaf@ugent.be

1b. Responsible Staff Member (ZAP)
------------------------------------------------------------
- name: Wouter Duyck
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: wouter.duyck@ugent.be
```

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies

===

* Reference of the publication in which the datasets are reported:

Dijkgraaf, Hartsuiker, and Duyck (Submitted). Modulating prediction of semantics in native and non-native speech comprehension: the role of cognitive load and processing speed

* Which datasets in that publication does this sheet apply to?:

All data sets reported in the publication

3. Information about the files that have been stored

===

3a. Raw data

----

* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

If NO, please justify:

* On which platform are the raw data stored?

  - [X] researcher PC
  - [ ] research group file server
  - [X] other (specify): external harddrive

* Who has direct access to the raw data (i.e., without intervention of another person)?

  - [X] main researcher
  - [ ] responsible ZAP
  - [ ] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files

----

* Which other files have been stored?

  - [ ] file(s) describing the transition from raw data to reported results. Specify:
  - [X] file(s) containing processed data. Specify: csv data files used for the analyses
  - [X] file(s) containing analyses. Specify: R script
  - [ ] files(s) containing information about informed consent
  - [ ] a file specifying legal and ethical provisions

- [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...

- [ ] other files. Specify: ...

* On which platform are these other files stored?

  - [X] individual PC

  - [ ] research group file server

  - [X] other (specify): external harddrive, and on open science framework: osf.io/8t76r

* Who has direct access to these other files (i.e., without intervention of another person)?

  - [X] main researcher

  - [ ] responsible ZAP

  - [ ] all members of the research group

  - [ ] all members of UGent

  - [X] other (specify): anyone

4. Reproduction

===

* Have the results been reproduced independently?: [ ] YES / [X] NO

* If yes, by whom (add if multiple):

   - name:

   - address:

   - affiliation:

   - e-mail:

v0.2

DATA STORAGE FACT SHEET CHAPTER 5

% Data Storage Fact Sheet

% Chapter 1

% Author: Aster Dijkgraaf

% Date: 28 June 2018

1. Contact details

===

1a. Main researcher

- name: Aster Dijkgraaf
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: aster.dijkgraaf@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Wouter Duyck
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: wouter.duyck@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies

===

* Reference of the publication in which the datasets are reported:

Broos, W.P.J., Dijkgraaf, A., Van Assche, E., Vander Beken, H., Dirix, N., Lagrou, E., Hartsuiker, R.J., Duyck, W. (in press). Is there adaptation of speech production after speech perception in bilingual interaction? Journal of Experimental Psychology: Learning, Memory, and Cognition.

* Which datasets in that publication does this sheet apply to?:

All data sets reported in the publication

3. Information about the files that have been stored

===

3a. Raw data

----

* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

If NO, please justify:

* On which platform are the raw data stored?
  - [X] researcher PC
  - [ ] research group file server
  - [X] other (specify): external harddrive

* Who has direct access to the raw data (i.e., without intervention of another person)?
  - [X] main researcher
  - [ ] responsible ZAP
  - [ ] all members of the research group

  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files

----

* Which other files have been stored?
  - [ ] file(s) describing the transition from raw data to reported results. Specify:
  - [X] file(s) containing processed data. Specify: csv data files used for the analyses
  - [X] file(s) containing analyses. Specify: R script
  - [ ] files(s) containing information about informed consent
  - [ ] a file specifying legal and ethical provisions
  - [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...
  - [ ] other files. Specify: ...

* On which platform are these other files stored?
  - [X] individual PC
  - [ ] research group file server
  - [X] other (specify): external harddrive, and on open science framework: osf.io/p62j4

* Who has direct access to these other files (i.e., without intervention of another person)?
  - [X] main researcher
  - [ ] responsible ZAP
  - [ ] all members of the research group
  - [ ] all members of UGent
  - [X] other (specify): anyone

4. Reproduction

===

* Have the results been reproduced independently?: [ ] YES / [X] NO

* If yes, by whom (add if multiple):
   - name:
   - address:
   - affiliation:
   - e-mail:

v0.2