Cross-linguistic influence during real-time sentence processing in bilingual children and adults

Published by
LOT
Kloveniersburgwal 48
1012 CX Amsterdam
The Netherlands

phone: +31 20 525 2461

e-mail: lot@uva.nl
http://www.lotschool.nl

ISBN: 978-94-6093-383-7
DOI: https://dx.medra.org/10.48273/LOT0598
NUR: 616

Cross-linguistic influence during real-time sentence processing in bilingual children and adults

Proefschrift
ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op vrijdag 17 september 2021
om 10.30 uur precies

door
Chantalle Nicole van Dijk
geboren op 26 september 1988
te Apeldoorn

Promotor:
Prof. dr. A.F.J. Dijkstra

Copromotor:
Dr. S. Unsworth

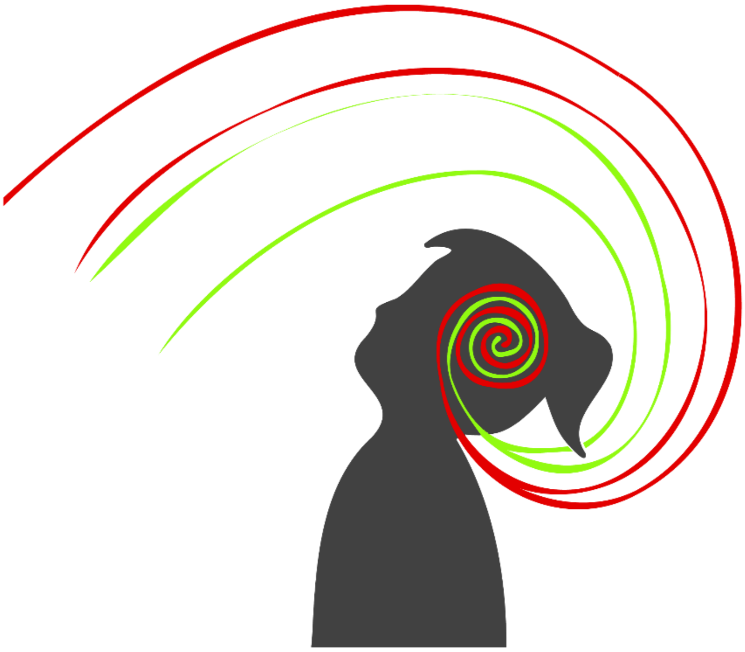
Manuscriptcommissie:
Prof. dr. R. Schoonen
Prof. dr. L. Serratrice (University of Reading, Verenigd Koninkrijk)
Prof. dr. R.J. Hartsuiker (Universiteit Gent, België)
Prof. dr. T. Marinis (Universität Konstanz, Duitsland)
Dr. J. Verhagen (Universiteit van Amsterdam)

This thesis is part of the research programme *Promoties in de geesteswetenschappen* and is financed by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

Table of contents

CHAPTER 1

Introduction



1.1 Introduction

When bilingual children speak in one of their languages, you might not even realise that they know another language. However, sometimes they say things in a slightly different way than a monolingual peer would do under influence of their other language. For example, a French-Dutch bilingual child might ask her brother *Waarom je huilt?* (why you cry; taken from Strik & Pérez-Leroux, 2011, p. 194). In French, this word order is correct. In Dutch, however, the subject and verb should be inverted: *Waarom huil je?* (why cry you, “Why are you crying?”). The use of French word order in this way while speaking Dutch is an example of *cross-linguistic influence*, the topic of this thesis.

Whilst cross-linguistic influence is well-studied in children’s speech production (e.g., Foroodi-Nejad & Paradis, 2009; Müller & Hulk, 2001; Strik & Pérez-Leroux, 2011), little is known about how morphosyntactic rules in one language influence processing of the other language while children *listen* to speech. Imagine, for instance, a Turkish-Dutch bilingual child listening to the following sentence in Dutch:

Anna vertelde Sophie gisteren wat zij voor haar verjaardag krijgt.
Anna told Sophie yesterday what she for her birthday gets
“Anna told Sophie yesterday what she would get for her birthday.”

In order to understand the sentence, the child must complete a number of tasks. First of all, she has to segment the speech wave and convert it into spoken words. Then, she has to search for the meaning of the words in her mental lexicon. However, knowing the meaning of the separate words is not enough to interpret the sentence. To be able to do this, the child has to build a syntactic structure to connect the different words and assign them their thematic roles. Still, this is not sufficient to completely interpret the sentence. At this moment in processing, the child might know that someone gets something for her birthday, but she does not know yet who that someone is. In other words, the child has to link the personal pronoun *zij* (as well as the possessive pronoun *haar*) to an antecedent in the discourse, which can either be Anna or Sophie. In Dutch, the preference would be for the topic of the discourse, Anna.

Each of these steps needs to occur rapidly and in an incremental fashion for successful comprehension to take place (e.g., Altmann & Steedman, 1988; Levelt, 1989). This means that the child must have acquired all the relevant linguistic knowledge in Dutch, such as the words used in the sentence, the syntactic structures, and the topic-preference for Dutch pronouns. Moreover, the child has to be able to apply her linguistic knowledge in a short amount of time while keeping incoming information about the sentence in working memory. The Turkish-Dutch bilingual child in our example will also have knowledge of Turkish. Hence, during sentence processing in Dutch, she also has to be able to select the knowledge that is relevant for Dutch and ignore information from Turkish. What happens, however, when the morphosyntactic rules of her languages differ? For instance, in the Turkish equivalent of our example above, the pronoun refers back to Sophie rather than to Anna. Will this difference between languages influence the child's interpretation of the Dutch sentence during listening? In other words, do rules from one language influence the processing of another language in bilingual children? This is the main question to be answered in this thesis.

Cross-linguistic influence during sentence processing in bilingual children

Recent models of speech production in bilingual children suggest that cross-linguistic influence can indeed take place during sentence processing (e.g., Nicoladis, 2006, 2012; Nicoladis, Rose, & Foursha-Stevenson, 2010; Serratrice, 2007, 2016; Sorace & Serratrice, 2009). Such models are based on the adult (L2) literature (e.g., Costa, 2004; Dell, Chang, & Griffin, 1999; Ferreira & Dell, 2000; Hartsuiker, Pickering, & Veltkamp, 2004; Hartsuiker & Bernolet, 2017) and assume that language co-activation at the lexical level results in language co-activation at the syntactic level. Figure 1.1 shows an example of such a model for adjective-noun orders in English and French (adapted from Nicoladis, 2006).

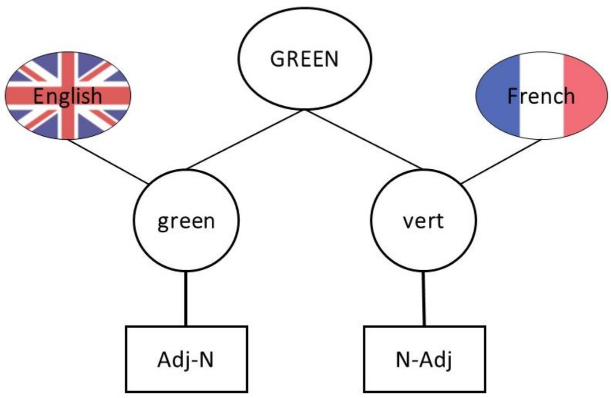


Figure 1.1. Example of lexical entries for *green* and *vert* in Hartsuiker et al.'s (2004) activation model in bilinguals with each lemma node (e.g., *green* and *vert*) connected to separate combinatorial nodes. GREEN represents the semantic representation of the lemmas *green* and *vert*. The flags refer to the language membership of the lemmas (Dijkstra & van Heuven, 2002).

In English, attributive adjectives are prenominal (Adj-N; e.g., *green apple*). In French, most attributive adjectives are postnominal (N-Adj; e.g., *pomme verte*, “apple green”), but some are prenominal (Adj-N; e.g., *petite pomme*, “small apple”). In this model, when bilingual French-English children produce an adjective-noun structure (e.g., *pomme verte*) in French, this activates both the French and the English lemmas (i.e, *vert* and *green*) through their semantic representation at the conceptual level (GREEN). In turn, the lemmas activate the kind of grammatical constructions a word can be used in, so-called combinatorial nodes (e.g., Schoonbaert, Hartsuiker, & Pickering, 2007): N-Adj in French and Adj-N in English. As a consequence, word orders from both languages compete for selection during speech production (e.g., Nicoladis, 2006, 2012; Nicoladis et al., 2010). Moreover, it has been argued for bilingual adults that with developing proficiency, similar syntactic representations in different languages, such as the prenominal Adj-N representations in English and French, ultimately become shared (e.g., Hartsuiker et al., 2004; Hartsuiker & Bernolet, 2017; schematically represented in Figure 1.2, adapted from Hartsuiker & Bernolet, 2017).

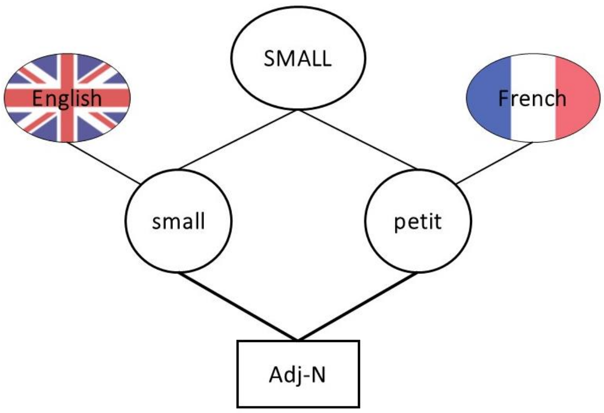


Figure 1.2. Example of lexical entries for *green* and *vert* in Hartsuiker et al.’s (2004) activation model in bilinguals with each lemma node (e.g., *small* and *petit*) connected to the same combinatorial node. SMALL represents the semantic representation of the lemmas *small* and *petit*. The flags refer to the language membership of the lemmas (Dijkstra & van Heuven, 2002).

There is some evidence for language co-activation in bilingual children that supports the models given in Figure 1.1 and Figure 1.2. First, evidence for lexical co-activation in bilingual children comes from studies on priming during lexical processing (e.g., Poarch & van Hell, 2012; Von Holzen, Fennell, & Mani, 2019; Von Holzen & Mani, 2012). For example, Von Holzen and Mani (2012) tested German-English bilingual children's recognition of words through an adaptation of the preferential looking paradigm. Children were found to recognise words in English faster (e.g., *slide*) when they were preceded by a phonologically similar prime in German (e.g., *Kleid*, "dress") than when they were preceded by an unrelated word in German. Crucially, the speed at which children recognized a word was reduced when that word (e.g., *Stein*, "stone") followed a prime that was phonologically related through its translation (e.g., *leg* through its German translation *Bein*). These findings show that during sentence processing words in a bilingual child's one language can activate related words in her other language.

Second, there is also evidence for co-activation at the syntactic level in bilingual children (e.g., Hsin, Legendre, & Omaki, 2013; Vasilyeva et al., 2010). For example, Vasilyeva and colleagues (2010) tested Spanish-English children's production of passive sentence structures. They found that children were more likely to produce a passive structure in English when they had just heard a passive structure in Spanish rather than an active structure. Furthermore, Hsin and colleagues (2013) observed that listening to prenominal adjective orders in English (e.g., *a green apple*) could prime the use of the same, but ungrammatical, order in Spanish in English-Spanish children (e.g., **una verde manzana*, "a green apple"). Hence, findings from these studies suggest that a structure in bilingual children's one language can activate and prime similar structures in their other language.

Given the observations that bilingual children's languages activate each other at both the lexical and the syntactic level, it seems rather unlikely that sentence processing in one language is completely unaffected by children's other language. Indeed, effects of online cross-linguistic influence due to language co-activation and priming have been observed in adults (discussed in more detail below; e.g., Hopp, 2017; Runnqvist, Gollan, Costa, & Ferreira, 2013). For instance, in a study targeting the same structure as our pronoun example at the start of this chapter, Schimke et al. (Schimke, de la Fuente, Hemforth, & Colonna, 2018) found that pronoun interpretation during the processing of a second language (L2) in adults was affected by properties of their first language (L1). Direct evidence for cross-linguistic influence in real-time sentence processing in bilingual children is lacking, however.

Surprisingly, hardly any study has investigated (morpho)syntactic activation across languages directly during sentence comprehension in bilingual children (but cf. Lemmerth & Hopp, 2019). Instead, studies that have examined language processing in this population have typically focused on the more general question of whether bilingual children can process morphosyntactic properties in a qualitatively similar manner as monolingual children (e.g., Blom & Vasić, 2011; Chondrogianni & Marinis, 2012; Chondrogianni, Marinis, Edwards, & Blom, 2015; Chondrogianni, Vasić, Marinis, & Blom, 2015; Marinis, 2007; Marinis & Saddy, 2013). Such studies employed so-called online tasks, such as the self-paced listening, word monitoring, or visual world paradigm (e.g., Blom & Vasić, 2011; Chondrogianni & Marinis, 2012; Chondrogianni, Marinis, et al., 2015; Chondrogianni, Vasić, et al., 2015; Lemmerth & Hopp, 2019). For example, Chondrogianni and Marinis (2012) investigated Turkish-English children's production of the third person marker *-s* and past tense maker *-ed* and their sensitivity to the omission of these tense markers. Bilingual children omitted these tense markers more often than their monolingual English peers. However, they showed online sensitivity to tense markers in a word-monitoring task. More specifically, when tense was incorrectly omitted, children were slower to perform the task, similar to monolingual children.

Results from such online studies could shed some light on the question whether cross-linguistic influence occurs during real-time sentence processing, even though investigating cross-linguistic influence was not the goal of the studies. This is because most studies investigated the processing of morphosyntactic properties in bilingual children's one language that were present or absent in their other language. For instance, both of the bilingual children's languages in Chondrogianni and Marinis' (2012) study (i.e., English and Turkish) use tense marking. Hence, the observation that bilingual children were as sensitive to tense marking online as English monolingual peers could be a facilitative effect of their Turkish. In the same vein, the observation by Blom and Vasić (2011) that Turkish-Dutch children were less sensitive than monolingual peers to gender agreement violations between nouns and their determiners (e.g., **de$_{COM}$ paard$_{NEUT}$*, "the$_{\text{COM}}$ horse$_{\text{NEUT}}$") could be accounted for by the absence of grammatical gender in Turkish.

At the same time, it is impossible to unambiguously contribute similarities and differences between bilingual and monolingual children's performance in these studies to cross-linguistic influence, because no bilingual control groups were involved. Consequently, Turkish-Dutch children's sensitivity to tense marking in English might have unfolded online irrespective of their knowledge of tense marking in Turkish, simply because they had acquired this aspect of English. Similarly, Turkish-Dutch children's insensitivity to gender violations online in Dutch could be a consequence of insufficient exposure to Dutch rather than the absence of grammatical gender in Turkish. The latter is not unlikely given that in almost all of these online studies, including Blom and Vasić (2011), children were tested in a second language that they were still in the process of acquiring (but again, cf. Lemmerth & Hopp, 2019). As a consequence, existing online studies cannot unequivocally determine whether the morphosyntactic properties of one language can influence bilingual children's online comprehension of another language.

Why offline research is not enough

In contrast to online studies, cross-linguistic influence has been the subject of investigation in children's offline sentence comprehension (e.g., Kidd, Chan, & Chiu, 2015; Nicoladis, 2002, 2003; Serratrice, 2007), for instance, using picture selection tasks. Although these provide some evidence for cross-linguistic influence (e.g., Kidd et al., 2015; Serratrice, 2007), it is unclear to what extent the outcomes of such tasks accurately reflect cross-linguistic influence during real-time sentence comprehension.

First of all, offline comprehension tasks involve a strong working memory component (Marinis, 2010). Indeed, not only do children have to process a sentence, they also have to keep it in working memory while making some kind of decision – i.e., choose the corresponding picture. Given that having to deal with two languages at the same time might already put a processing burden on bilingual children (e.g., Sorace, 2011; Sorace & Serratrice, 2009), the additional working memory component of offline tasks might affect their performance more strongly than that of monolingual children.

In addition, offline tasks allow children to use their explicit language knowledge and meta-linguistic abilities to inform their responses (Marinis, 2010). These explicit strategies might override children's online decisions and preferences. Furthermore, if cross-linguistic influence during real-time sentence comprehension is subtle, it might not be strong enough to surface in an offline task. Take, for instance, the processing of the pronoun *zij* ("she") in the example at the start of this chapter. A Turkish-Dutch child might consider *Sophie* as an alternative option to *Anna* to a greater extent than monolingual peers during real-time comprehension. However, the preferred Dutch interpretation *Anna* might be so strong that co-activation of the Turkish discourse-pragmatic strategy does not influence the child's final interpretation. Consequently, the child's interpretation of the pronoun *after* processing has taken place, as measured by an offline comprehension task, might not reflect a subtle online interplay of Turkish and Dutch.

In sum, at present hardly anything is known about cross-linguistic influence during real-time sentence comprehension in bilingual children at the level of morphosyntax. Evidence from existing online studies is inconclusive. Furthermore, evidence from offline studies concerning cross-linguistic influence in comprehension is limited and indirect at best. Therefore, the main aim of this thesis is to employ online techniques to assess cross-linguistic influence during real-time sentence processing in bilingual children.

In order to accomplish this aim, we combine insights from the offline child bilingualism literature with insights from the online adult second language (L2) literature. Previous studies with bilingual children have identified a number of predictors of cross-linguistic influence, surface overlap, language dominance and age. These predictors might be relevant for online cross-linguistic influence as well. Cross-linguistic influence during sentence processing is a central and well-studied theme within the field of adult L2 acquisition. Hence, research techniques and observations from this field can inform studies with bilingual children.

To further clarify the objectives of this thesis, we first present an overview of results from production and offline comprehension studies on cross-linguistic influence in bilingual children in section 2. In section 3 we discuss studies on online cross-linguistic influence in adult second language learners. In section 4, we present the research questions of this thesis and outline the individual chapters that follow.

1.2 Cross-linguistic influence in production and offline comprehension in bilingual children

In the past three decades, morphosyntactic development has been one of the key areas of research on bilingual children and it continues to receive considerable attention (e.g., Guijarro-Fuentes, 2019; Serratrice, 2013). Findings have resulted in the perhaps at first sight paradoxical claim that whilst bilingual children are well able to differentiate between their language systems from early on (e.g., Houwer, 1990; Meisel, 2011; Paradis & Genesee, 1996), morphosyntactic properties from one language can influence the other language (e.g., Hulk & Müller, 2000; Serratrice, 2013). Cross-linguistic influence can result in qualitative differences between bilingual and monolingual children (e.g., Döpke, 1998; Strik & Pérez-Leroux, 2011; Yip & Matthews, 2000). For instance, the lack of subject-verb inversion in *wh*-questions in the French-Dutch bilingual child in the example at the start of this chapter (*Waarom je huilt?*, “Why are you crying?”) has not been observed in monolingual acquisition (e.g., Strik and Pérez-Leroux, 2011).

A more common observation is quantitative cross-linguistic influence (e.g., Argyri & Sorace, 2007; Foroodi-Nejad & Paradis, 2009; Nicoladis & Gavrila, 2015). This has manifested itself as a difference in the frequency with which a certain structure is accepted or used by bilingual and monolingual children (Yip & Matthews, 2000). For example, Nicoladis and Gavrila (2015) found that Welsh-English bilingual and English monolingual children sometimes produced ungrammatical postnominal adjectives in English (e.g., **apple green* instead of *green apple*). The bilingual children did so to a greater extent than the monolingual children. The authors attributed the difference between the groups to cross-linguistic influence from the Welsh postnominal adjective order into English.

Cross-linguistic influence has been observed for various morphosyntactic properties and in various language combinations (see Serratrice, 2013 for an overview). However, whether or not cross-linguistic influence is present and how strong its effect is, differs from study to study. A number of variables that can account for such variation have been put forward, such as surface overlap between languages, the (morpho)syntactic domain investigated, language dominance, the quality of input children receive, and children's chronological age (e.g., Argyri & Sorace, 2007; Bosch & Unsworth, 2020; Döpke, 1998; Hulk & Müller, 2000; Paradis & Navarro, 2003; Sorace & Serratrice, 2009). In this thesis, we focus on the role of three predictors – surface overlap, language dominance and age – discussed in more detail below.

Surface overlap

With regard to surface overlap, Hulk and Müller (2000) originally proposed that there has to be a certain ambiguity in the child's language input in order for cross-linguistic influence to occur. More in particular, if a certain structure in language a can be analysed by syntactic analysis X or Y and language α provides evidence for analysis X only, language α may reinforce the use of that analysis in language a, resulting in cross-linguistic influence (Hulk & Müller, 2000; Müller & Hulk, 2001, also see Döpke, 1998, for a similar account).

To illustrate a situation of surface overlap, we turn to compounding in English and Persian (Foroodi-Nejad & Paradis, 2009). In English, compounds are – with a few exceptions – right headed (e.g., apple *juice*). In Persian, in contrast, compounds are preferably left-headed, but can be right-headed as well (see Table 1.1). The availability of two options for compounding in Persian might lead monolingual and bilingual children to conclude that right-headed compounds are possible in situations where they are not allowed by the adult language. Furthermore, the availability of right-headed compounds only in

English might reinforce the incorrect or dispreferred right-headed compounds in Persian for bilingual children even more, in line with the surface overlap hypothesis, as indicated by the arrow in Table 1.1. This is indeed what was observed by Foroodi-Nejad and Paradis (2009) in their study of bilingual Persian-English children's acquisition of compounding.

Table 1.1. Compound order in English and Persian.

English		Persian
Right-headed *apple juice*	→	Right-headed $[ab_N\ sib_N]_N$ water apple ("apple juice")
		Left-headed $[gol_N\ ab_N]_N$ flower water ("flower juice")

On the whole, however, evidence for effects of surface overlap is mixed. Many studies have found evidence for cross-linguistic influence in situations of surface overlap (e.g., Foroodi-Nejad & Paradis, 2009; Hacohen & Schaeffer, 2007; Haznedar, 2007; Mykhaylyk & Ytterstad, 2017; Schmitz, Patuto, & Müller, 2011), and at the same time not in situations without surface overlap (e.g., Austin, 2009; Müller & Hulk, 2001; Serratrice, Sorace, Filiaci, & Baldo, 2009). However, differences between bilingual and monolingual children have also been shown in situations *without* overlap (e.g., Argyri & Sorace, 2007; Foroodi-Nejad & Paradis, 2009; Nicoladis, 2006).

Language dominance

The second predictor of interest investigated in this thesis is language dominance. Dominance can be operationalized in various ways, such as the amount of exposure children receive and their relative proficiency in their languages (e.g., Silva-Corvalan & Treffers-Daller, 2016). In theory, it would be possible for a child to be exactly balanced in their two languages – i.e., receive an equal amount of exposure and be equally proficient in language *a* and *α*. However, in practice, children typically are more dominant in one of their languages. Furthermore, dominance is not a static construct within a bilingual child. Instead, children's dominant language can shift during various stages of life (e.g., De Houwer, 2011; Meisel, 2007; Yip & Matthews, 2000).

With regard to cross-linguistic influence, the observation is that children's dominant language is more likely to influence their non-dominant language than vice versa (e.g., Argyri & Sorace, 2007; Foroodi-Nejad & Paradis, 2009; Yip & Matthews, 2000). In some studies, language dominance dictated the direction of cross-linguistic influence (e.g., Argyri & Sorace, 2007; Sorace, Serratrice, Filiaci, & Baldo, 2009; Yip & Matthews, 2000). Cross-linguistic influence was only observed in the direction of the dominant language into the non-dominant language. In other studies, language dominance was found to affect the strength of cross-linguistic influence, rather than its occurrence (e.g., Bosch & Unsworth, 2020; Foroodi-Nejad & Paradis, 2009; Kidd et al., 2015). To be more precise, their observation was that cross-linguistic influence was stronger from the dominant into the non-dominant language than vice versa. However, there are also studies that found no relationship between language dominance and the occurrence or strength of cross-linguistic influence (e.g., Foroodi-Nejad & Paradis, 2009; Nicoladis, 2002; Unsworth, 2012).

Age

Our final predictor of interest is children's chronological age. In earlier studies with young bilingual children, it is generally assumed that cross-linguistic influence is a temporary phenomenon. Specifically, cross-linguistic influence has been argued to facilitate or delay morphosyntactic acquisition (e.g., Hulk & Müller, 2000; Paradis & Genesee, 1996). However, evidence for the role of age from experimental studies is mixed. On the one hand, studies have shown cross-linguistic influence in older bilingual children (e.g., Argyri & Sorace, 2007; Daskalaki, Chondrogianni, Blom, Argyri, & Paradis, 2019; Serratrice et al., 2009; Sorace et al., 2009). For example, Argyri and Sorace (2007) observed cross-linguistic influence in a group of children between 7;5 and 9;5 years old. Furthermore, a number of studies found no significant relationship between children's age and the observed effect of cross-linguistic influence (e.g., Bosch & Unsworth, 2020; Nicoladis, 2002; 2003). These findings suggest that cross-linguistic influence can still occur after language acquisition has taken place. On the other hand, some studies observed the effect of cross-linguistic influence to diminish or disappear completely with age (e.g., Serratrice et al., 2009; Sorace et al., 2009; Unsworth, 2012). Furthermore, cross-linguistic influence has even been found to increase with age (e.g., Nicoladis & Gavrila, 2015). As a consequence, the effect of age on cross-linguistic influence is as yet poorly understood.

Studies with simultaneous bilingual adults could inform the debate on whether cross-linguistic influence persists once language acquisition is complete (e.g., Kupisch, 2012, 2014; Kupisch & Barton, 2013; Runnqvist et al., 2013; Schmitz, Di Venanzio, & Scherger, 2016). Indeed, there is some evidence that cross-linguistic influence occurs in adults, but mainly from the dominant into the non-dominant language (e.g., Kupisch, 2012, 2014).

However, very few (online) studies exist investigating this population, such that evidence is scarce. Moreover, research designs typically differ between child and adult studies, making a direct comparison between populations difficult. As a result, it also remains unknown to what extent cross-linguistic influence observed in (young) bilingual children remains present after acquisition has taken place.

To sum up, although cross-linguistic influence is a frequently studied phenomenon in bilingual children, we still understand little about the circumstances it appears in and, consequently, the mechanisms behind it. Furthermore, the majority of studies have focussed on cross-linguistic influence in bilingual children's sentence production rather than in sentence comprehension. Moreover, as argued before, there are virtually no studies on cross-linguistic influence during sentence processing in simultaneous bilingual children (but cf. Lemmerth & Hopp, 2019) and very few in simultaneous bilingual adults (e.g., Martohardjono, Phillips, Madsen II, & Schwartz, 2017). As a consequence, it is as yet unknown how surface overlap and language dominance affect cross-linguistic influence in children online and whether cross-linguistic influence persists with age.

1.3 Cross-linguistic influence during sentence processing in adult L2 learners

In contrast to child studies, cross-linguistic influence during real-time sentence comprehension in adult L2 learners has been studied for a wide range of language properties, such as relative clause attachment, pronoun resolution and word order. Frequently used methodologies include self-paced reading, eye-tracking and ERPs (e.g., Alemán Bañón, Fiorentino, & Gabriele, 2018; Cunnings, Fotiadou, & Tsimpli, 2017; Foucart & Frenck-Mestre, 2012; Hopp, 2010; Schimke et al., 2018). Findings from such studies can be assigned into three categories, which are relevant to this thesis: (i) native-like; (ii) affected by general processing difficulties; (iii) affected by their L1.

Native-like L2 processing

With sufficient proficiency, L2 learners have been found to process their L2 in qualitatively similar ways as native speakers, even when the property being studied was absent or different in participants' L1 (e.g., Alemán Bañón et al., 2018; Gillon Dowens, Guo, Guo, Barber, & Carreiras, 2011; Witzel, Witzel, & Nicol, 2012). For example, Alemán Bañón and colleagues (2018) investigated online sensitivity to number- and gender-agreement violations in Spanish in L2 learners with English as L1. They did so by measuring participants' ERPs and by comparing these to ERPs in native speakers of Spanish. Spanish adjectives have to agree both in number and gender to the noun they modify (e.g., *fruta jugosa/*fruit$_{FEM\text{-}SING}$ juicy$_{FEM\text{-}SING}$, "juicy fruit"; adapted from Alemán Bañón et al., 2018, p. 10). Whilst English has number agreement, it is not instantiated on adjectives. Furthermore, English has no gender agreement.

The authors found that the lowest proficiency participants only showed sensitivity to number violations, as evidenced by P600 effect, and not to gender violations. At the same time, the highest proficient L2 learners showed robust P600 effects for both type of violations, similar to native speakers of Spanish. Hence, regardless of the different properties of English and Spanish, L2 learners were able to acquire nativelike processing abilities of number and gender violations.

General L2 processing difficulties

Although native-like processing behaviour is possible in L2 learners, the vast majority of online studies show that L2 learners behave different from native speakers (see e.g., Cunnings, 2017; Kroll & Dussias, 2013, for an overview). Most of these studies account for these results in terms of general processing difficulties. To be more precise, L2 speakers are argued to be less efficient in processing their second language due to general differences between L2 learners and native speakers (e.g., Cunnings et al., 2017; Felser, Roberts, Marinis, & Gross, 2003; Hopp, 2010; Papadopoulou & Clahsen, 2003; Roberts, Gullberg, & Indefrey, 2008).

For instance, Roberts and colleagues (2008) studied Dutch pronoun resolution in Dutch native speakers and L2 learners with either Turkish or German as L1. In German, like in Dutch, the pronoun usually refers to the topic of the discourse, such as *Anna* in our example at the start of this chapter. In contrast, pronouns in Turkish usually signal a shift in topic. In an eye-tracking during reading task, Roberts and colleagues found Turkish participants to experience difficulties interpreting Dutch pronouns, as evidenced by slowdown effects. A similar pattern was observed for the L1 German group, but not in Dutch native speakers. Therefore, the authors argued, L2 learners had difficulties integrating discourse information during syntactic processing, regardless of the properties of their L1.

The central question with regard to the observed general processing differences between L2 learners and native speakers is to what extent these are quantitative or qualitative in nature. Some have argued that such differences are caused by qualitative differences in the processing mechanism of L2 learners (e.g., Clahsen & Felser, 2006). In particular, Clahsen and Felser (2006) proposed that L2 learners only construct shallow syntactic structures without much syntactic detail. Instead, they assumed that L2 learners rely more on semantics and pragmatics than on syntactic cues.

Others have attributed L2 effects to quantitative differences in processing between L2 learners and native speakers (e.g., Cunnings, 2017; Hopp, 2010; Sorace, 2011). For instance, Sorace (2011) has proposed that L2 learners have difficulties integrating different sources of information during processing, in particular discourse-pragmatics. This, she argued could be due to insufficient processing resources in L2 learners compared to native speakers. A proposal along the same lines has been put forward by Hopp (2010). He noticed the presence of similar processing difficulties during morphosyntactic processing in German native speakers as in L2 learners when processing demands were sufficiently high. Hence, Hopp's findings offer support for a 'quantitative-difference' view between L2 learners and native speakers, rather than a qualitative one.

An important observation is therefore that language processing behaviour in L2 learners can deviate from language processing behaviour in native speakers due to general L2 effects, and not necessarily due to cross-linguistic influence. Discussing the exact mechanisms that have been proposed to account for such differences between native and L2 processing in more detail is outside of the scope of this thesis, however (for more information on this topic, see e.g., Clahsen & Felser, 2006; Cunnings, 2017; Hopp, 2010; Kroll & Dussias, 2013; Sorace, 2011).

Cross-linguistic influence in L2 processing

Finally, findings of a number of studies with adult L2 learners can be classified as cross-linguistic influence from the L1 into the L2 (e.g., Foucart & Frenck-Mestre, 2012; Hopp, 2017; Runnqvist et al., 2013; Schimke et al., 2018). For example, Hopp (2017) tested for effects of cross-linguistic influence during sentence reading. He investigated reading speed in L2 learners of English with German as L1 and in native speakers of English. The word order of sentences tested either did not overlap at all with German (e.g., *The doctor Sarah ignored tried to leave the room (…),* p. 105) or overlapped but differed in meaning (e.g., *When the doctor Sarah ignored tried to leave the room (…),* p. 105). Hopp found low proficient L2 learners of English to slow down when reading the latter type of sentences compared to native speakers. He observed similar effects in highly proficient L2 learners when English stimuli were alternated with German sentences (bilingual mode). When there was no overlap between sentences, L2 learners and native speakers showed similar patterns. Hence, Hopp's findings suggest that the overlapping word order in German affected the processing of the similar word order in English. Effects of overlapping word order from the L1 on the L2 have been found in L2 learners' ERPs and speed of production as well (e.g., Foucart & Frenck-Mestre, 2012; Runnqvist et al., 2013).

Further evidence for cross-linguistic influence in adult L2 learners comes from Schimke et al.'s (2018) study on pronoun resolution. In this study, online pronoun comprehension was tested in German by means of the visual world paradigm. Participants were L2 learners of German with either Spanish or French as L1. Participants listened to sentences with a pronoun that could either refer to the subject or the object of the previous sentence (e.g., *Der Straßenfeger$_i$ ist dem Briefträger$_k$ begegnet, bevor er$_{i/k}$ sehr schnell die Briefe geholt hat*, "The street sweeper$_i$ met the postman$_k$ before he$_{i/k}$ quickly fetched the letters", p. 765). In German and French, there is a strong preference for a pronoun to be linked to the subject of the previous sentence ("the street sweeper"), whereas in Spanish the pronoun is either linked to the object ("the postman") or to neither the subject or the object. Schimke and colleagues found that after hearing the pronoun L1 French learners of German looked more to the picture of the subject (the street sweeper) than of the object (the postman). This was in line with patterns observed for German native speakers. In contrast, the L1 Spanish group showed no preference for the subject or object interpretation, in line with cross-linguistic influence from Spanish.

Findings of cross-linguistic influence during sentence processing in adult L2 learners have been accounted for in terms of cross-language competition effects and priming (e.g., Hartsuiker & Bernolet, 2017; Hopp, 2010, 2017; Runnqvist et al., 2013). With regard to cross-language competition, Hopp (2017) argued that for the L2 learners in his study, processing the overlapping word order in English activated the same word order in German. In turn, they had to allocate processing resources to inhibit this competition from German. Consequently, less processing resources were available to interpret the sentence. As a result, L2 learners sometimes resorted to constructing a sentence structure in line with their L1 German rather than their L2 English syntax. This caused slowdown effects during reading. Hence, Hopp (2017) directly linked accounts of less efficient processing in L2 learners to L1 effects (e.g., Hopp, 2010; Sorace, 2011).

A different account is put forward by Runnqvist and colleagues (2013). They discussed online cross-linguistic influence in a framework of shared syntactic structures, namely based on the proposal that L2 learners develop one shared syntactic representation when syntactic structures overlap between their L1 and L2 (e.g., Hartsuiker et al., 2004; Hartsuiker & Bernolet, 2017). Consequently, when a shared syntactic structure is more frequent in L2 learners' L1 or L2, it is more readily available in their other language as the result of cross-language priming (e.g., Runnqvist et al., 2013). Hence, existing theories on cross-linguistic influence in bilingual adults suggest that the processing of a structure in one language might be facilitated or delayed by the availability of the same structure in the other language.

To sum up, sentence processing in L2 learners can be qualitatively similar to sentence processing in native speakers or differ due to either general bilingualism effects or cross-linguistic influence from the L1. Previous online studies with bilingual children suggest that bilingual and monolingual children can process language in qualitatively similar ways as well (e.g., Chondrogianni & Marinis, 2012; Marinis, 2007; Marinis & Saddy, 2013). However, it is as yet unclear whether cross-linguistic influence and general bilingualism effects can affect sentence processing in bilingual children as well.

With regard to cross-linguistic influence, online effects can be expected in bilingual children for two reasons. First of all, as discussed above, mechanisms that have been argued to underlie cross-linguistic influence during sentence processing in L2 adults – i.e., language co-activation and priming – have also been observed in bilingual children. Hence, these may result in online cross-linguistic influence in children as well (also see, e.g., Nicoladis, 2006, 2012; Nicoladis et al., 2010; Serratrice, 2007, 2016; Sorace & Serratrice, 2009). Second, there is ample evidence for cross-linguistic influence for various morphosyntactic properties from offline comprehension and production studies in bilingual children (e.g., Serratrice, 2013). It is possible that such cross-linguistic influence reflects cross-linguistic influence during real-time sentence processing.

In addition, bilingual children's online sentence comprehension might be influenced by general processing difficulties, similar to L2 adults. Again, there are at least two reasons why we might expect such effects in children. First, bilingual children are likely to receive relatively less input in their two languages than monolingual children in their one language. Consequently, bilingual children will have less experience processing their individual languages than monolingual peers. Therefore, their processing might be less automatized and, as a result, less efficient (e.g., Sorace, 2011; Sorace & Serratrice, 2009). Second, due to bilingual children's languages being always co-activated, children have to allocate processing resources to inhibit activation of the language not in use. As a consequence, they might have insufficient processing resources available for efficient language processing (e.g., Hopp, 2017; Sorace, 2011; Sorace & Serratrice, 2009).

At the same time, findings from adult L2 learners should not automatically be extended to bilingual children. That is because bilingual language development in children obviously differs from adult L2 acquisition in the sense that simultaneous and early sequential bilingual children – the object of inquiry in this thesis – acquire two native languages in parallel, rather than one after the other has been completely acquired. Consequently, from the start of language acquisition, children have been trained in differentiating their languages (e.g., Paradis & Genesee, 1996). In this light, then, the cross-linguistic influence and difficulties during sentence processing observed for adults in their L2 might not be experienced by children when processing their two languages.

1.4 Thesis outline

This thesis investigates whether and when cross-linguistic influence takes place during real-time sentence processing in bilingual children at the level of morphosyntax. In doing so, it combines insights from the child bilingualism and the adult L2 literature. With regard to the child bilingualism literature, we investigated whether predictors of cross-linguistic influence in children's offline comprehension and production affected cross-linguistic influence during real-time sentence processing as well. With regard to the adult L2 literature, we adapted online techniques previously employed to study cross-linguistic influence in adults for use with children. More in particular, we adopted the self-paced listening paradigm (Chapters 3 and 4) and eye-tracking in the visual world paradigm (Chapter 5). Additional aims were to compare offline and online findings of cross-linguistic influence and to test for general bilingualism effects during processing. Crucially, in order to distinguish effects of cross-linguistic influence from general bilingualism effects, we always compared results from two bilingual groups with different language combinations to a group of monolingual peers.

Our research questions were as follows:

- To what extent and in what manner does cross-linguistic influence manifest itself during sentence processing in bilingual children? (Chapters 3 and 5)
- To what extent is online cross-linguistic influence predicted by surface overlap between languages and by language dominance? (Chapters 3, 4 and 5)
- Is there evidence for general processing difficulties in bilingual children's sentence processing? (Chapters 3 and 5[1])

---

[1] Note that we do not address this question explicitly in Chapter 3. However, we do compare processing speed between bilingual and monolingual children in this chapter and discuss the results in relation to a general effect of bilingualism in the Discussion.

- How does cross-linguistic influence during sentence processing develop into adulthood? (Chapter 4)
- How do effects of cross-linguistic influence during real-time sentence processing relate to cross-linguistic influence in offline comprehension and production? (Chapters 2 and 5)

We now describe the objectives and research methodology of each chapter in more detail.

Chapter 2 consists of a literature review and meta-analysis of studies on cross-linguistic influence in bilingual children's production and offline comprehension. As discussed in the above, observed effects of cross-linguistic influence are variable from study to study. Furthermore, the role of our predictors of interest – surface overlap, language dominance and age – are not completely understood. Therefore, before conducting our empirical studies, we considered the construct of cross-linguistic influence and its predictors in more detail in this chapter. We did so in three steps: (i) in a systematic review we assessed how cross-linguistic influence and its predictors were operationalized in previous studies; (ii) in a meta-analysis we calculated the average weighted effect size and its consistency of cross-linguistic influence based on effects obtained in previous studies; and (iii) in separate meta-regressions we analysed effects of surface overlap, language dominance and age on cross-linguistic influence. Outcomes of Chapter 2 informed the next three empirical chapters on the operationalization of cross-linguistic influence, surface overlap and language dominance.

Chapters 3, 4 and 5 investigated cross-linguistic influence during sentence processing using online experimental techniques. The aim of Chapter 3 was twofold. First, we integrated findings from the child bilingualism and the adult L2 literature. In particular, in our literature review we examined whether observed offline effects of surface overlap and language dominance in bilingual children were present in online studies with adult L2 learners as well. We further used these insights to directly relate the role of surface overlap and language dominance in cross-linguistic influence in offline comprehension and production to accounts of language co-activation in bilingual children. Second, we investigated cross-linguistic influence at the level of word order from English to Dutch and German to Dutch by means of a self-paced listening task. Participants were 40 English-Dutch and 42 German-Dutch simultaneous bilingual children and 39 Dutch monolingual children. Word order in Dutch long passive and Verb Second sentences was systematically manipulated such that we could test for effects of surface overlap. Furthermore, we investigated effects of language dominance online. To our knowledge, this is the first study that investigated cross-linguistic influence during real-time processing of word order in bilingual children.

Chapter 4 is a direct follow-up of Chapter 3. The main aim of Chapter 4 was to investigate whether observed effects of cross-linguistic influence in Chapter 3 could be extended to simultaneous bilingual adults and adolescents. Most existing studies on cross-linguistic influence typically either focus on bilingual children using production and offline comprehension tasks, or on adult L2 learners, often using online tasks. Furthermore, as discussed in the above, the effect of age on the occurrence and strength of cross-linguistic influence in bilingual children is unclear. Consequently, little is known about how cross-linguistic influence in simultaneous bilingual children develops after acquisition has taken place, especially in online sentence processing. Therefore, Chapter 4 replicates the self-paced listening experiment from Chapter 3 with adults and adolescents with the same linguistic background as the children tested. In total, we tested 26 English-Dutch and 25 German-Dutch simultaneous bilingual adults and adolescents and 25 monolingually raised Dutch adults and adolescents. Again, we investigated the effect of surface overlap and language dominance on online cross-linguistic influence. In addition, we tested whether general bilingualism effects and language mode, as observed with adult L2 learners, affected the bilingual groups' behaviour during real-time sentence processing.

Chapter 5 investigated cross-linguistic influence during real-time pronoun resolution by means of the visual world paradigm. Pronoun resolution is extensively studied in bilingual children and adult L2 learners. However, only offline comprehension and production studies have been employed with bilingual children. Therefore, this study adapted online experiments from the field of adult L2 acquisition for use with children. We tested offline and online pronoun comprehension in Dutch in 17 Turkish-Dutch, 23 German-Dutch and 14 Dutch monolingual children. Furthermore, we assessed the effect of language dominance on children's offline and online behaviour. In addition, we tested for general bilingualism effects on children's processing patterns. To our knowledge, this is the first study to do so in bilingual children. Results of this study do not only shed light on the occurrence of cross-linguistic influence during sentence processing in bilingual children in general, but further inform the debate on cross-linguistic influence in the following two ways. First, this study was the first to directly compare effects of cross-linguistic influence in children's offline and online comprehension. Second, our set-up allowed us to compare online pronoun resolution in bilingual children to online pronoun resolution in adult L2 learners as observed in previous studies. In this way, we could examine whether similar mechanisms are at play during bilingual and L2 processing.

Finally, in **Chapter 6** we summarize the findings from Chapters 2 through 5. Furthermore, we discuss our results in relationship to existing literature on cross-linguistic influence in bilingual children and online cross-linguistic influence in adult L2 learners. We then integrate insights from our results and previous studies into a new model of online and offline cross-linguistic influence in bilingual children by extending existing accounts of language co-activation and priming. The new model is called CLISP, an acronym of Cross-Linguistic Influence during Sentence Processing.

CHAPTER 2

Cross-linguistic influence in simultaneous bilingual children

A meta-analysis

Abstract

Although cross-linguistic influence at the level of morphosyntax is one of the most intensively studied topics in child bilingualism, the circumstances under which it occurs remain unclear. In this meta-analysis, we measured the effect size of cross-linguistic influence and systematically assessed its predictors in 750 simultaneous and early sequential bilingual children in 17 unique language combinations across 26 experimental studies. We found a significant small to moderate average effect size of cross-linguistic influence, indicating that cross-linguistic influence is part and parcel of bilingual development. Language dominance, operationalized as societal language, was a significant predictor of cross-linguistic influence, whereas surface overlap, language domain and age were not. Perhaps an even more important finding was that definitions and operationalisations of cross-linguistic influence and its predictors varied considerably between studies. This could explain the absence of a comprehensive theory in the field. To solve this issue, we argue for a more uniform method of studying cross-linguistic influence.

Based on: van Dijk, C.N., van Wonderen, E., Koutamanis, E., Kootstra, G.J., Dijkstra, T., & Unsworth, S. (2021). Cross-linguistic influence in simultaneous bilingual children: a meta-analysis. *Journal of Child Language*. https://doi.org/10.1017/S0305000921000337

2.1 Introduction

How a bilingual child's two languages affect each other has been a prominent topic of research in the field of bilingual first language acquisition over the past three decades. Such *cross-linguistic influence*, most commonly investigated at the level of (morpho)syntax, has been attested in both the spontaneous and elicited speech production of simultaneous bilingual children, as well as in their comprehension and judgements of sentences (see Serratrice, 2013, for an overview). Cross-linguistic influence is defined here as the overuse or overacceptance of (morpho)syntactic properties in bilingual children's one language under influence of their other language. For example, Italian-English bilingual children have been found to overuse overt subject pronouns in Italian and this has been argued to result from cross-linguistic influence from English (e.g., Serratrice, Sorace, & Paoli, 2004). Researchers have aimed to identify the contexts in which cross-linguistic influence is most likely to appear. Well-studied predictors of cross-linguistic influence include surface overlap, language domain, language dominance, and age (e.g., Foroodi-Nejad & Paradis, 2009; Hulk & Müller, 2000; Müller & Hulk, 2001; Sorace & Serratrice, 2009; Yip & Matthews, 2000).

Evidence for the contribution of these predictors is mixed, however. Cross-linguistic influence is not always found when predicted (e.g., Argyri & Sorace, 2007; Nicoladis, 2002, 2003) and it is sometimes found when not predicted (e.g., Foroodi-Nejad & Paradis, 2009; Strik & Pérez-Leroux, 2011). Furthermore, cross-linguistic influence varies from child to child, as evidenced by the large standard deviations found in many studies (e.g., Mykhaylyk & Ytterstad, 2017; Nicoladis, 2006). As a consequence, there is neither consensus on the extent to which cross-linguistic influence in bilingual language acquisition takes place, nor what predicts it. To shed light on these issues, we conducted a meta-analysis to systematically examine the effect of morphosyntactic cross-linguistic influence in relation to surface overlap, language domain, language dominance, and age.

This paper is organised as follows. The next section discusses previous studies on cross-linguistic influence and the role of our predictors of interest. Then we list our research questions and hypotheses. The method section details our screening process, our coding procedure for surface overlap, language domain, language dominance and age, and how we calculated effect sizes for cross-linguistic influence. Subsequently, we present the outcomes of the meta-analysis and we discuss the results in relation to previous literature. Finally, we formulate recommendations for future studies based on our findings.

Morphosyntactic development in bilingual children

Research on cross-linguistic influence is embedded in a larger debate about the architecture of simultaneous bilingual children's language systems. In the pioneering work of the 1990s, researchers focussed on whether or not children's morphosyntactic systems developed independently from one another (e.g., Houwer, 1990; Meisel, 1989; Paradis & Genesee, 1996). Taking separate systems as a starting point, research in the subsequent two decades investigated the extent to which cross-linguistic influence occurred (e.g., Hulk & Müller, 2000; Meisel, 2007; Paradis & Genesee, 1996; Serratrice, 2013).

Early work on cross-linguistic influence considered young children's spontaneous speech production in (multiple) case studies. Researchers typically compared the development of morphosyntactic properties in bilingual and monolingual children over a period of time (e.g., Döpke, 1998; Hulk & Müller, 2000; Paradis & Genesee, 1996). On the one hand, bilingual children were found to behave in language-specific ways, showing that they were able to differentiate the morphosyntactic rules of their languages (e.g., Döpke, 1998; Paradis & Genesee, 1996). On the other hand, the two languages were found to influence each other in both quantitative and qualitative ways: quantitative when acquisition of a certain morphosyntactic property was facilitated or delayed in bilingual children under influence of their other language; and qualitative when bilingual children used a morphosyntactic property unattested in the speech of monolingual peers under influence of their other language (e.g., Müller & Hulk, 2001; Paradis & Genesee, 1996; Yip & Matthews, 2000).

More recent studies have typically employed experimental techniques, resulting in data on a wide range of linguistic properties and language combinations (see Serratrice, 2013 for an overview). These data have allowed researchers to systematically test for cross-linguistic influence under specific conditions in larger groups of bilingual children. Furthermore, they make it possible to study cross-linguistic influence not only on the basis of children's speech production, but also children's comprehension and judgements (e.g., Meroni, Smeets, & Unsworth, 2017; Serratrice, 2007). At the same time, the comparison between bilingual and monolingual peers has remained central. Experimental studies have found similar patterns of behaviour as those using spontaneous speech data: bilingual children differentiated between the morphosyntactic properties of their languages, but at the same time showed quantitative and – to a lesser degree – qualitative cross-linguistic influence (e.g., Argyri & Sorace, 2007; Nicoladis, 2006; Strik & Pérez-Leroux, 2011).

Some studies have investigated cross-linguistic influence by comparing different groups of bilingual children with each other rather than comparing bilinguals with monolinguals (e.g., Kaltsa, Tsimpli, & Argyri, 2019; Serratrice, Sorace, Filiaci, & Baldo, 2009; Sorace & Serratrice, 2009). Such a design allows researchers to manipulate morphosyntactic properties cross-linguistically whilst at the same time controlling for bilingual vs. monolingual status (and all that this may entail) – we return to this design in more detail in the discussion. Because the vast majority of (experimental) studies on cross-linguistic influence have used a monolingual control group alongside a single bilingual group, we have focussed on this design in the present study.

Despite the many studies on the topic, the circumstances under which cross-linguistic influence emerges remain elusive. Cross-linguistic influence has been attested in various language combinations, for different linguistic properties, and using different tasks, but findings are inconsistent. Study outcomes can differ even when the same morphosyntactic property in the same language was under investigation (compare Rodina et al., 2020; Schwartz et al., 2015). Various predictors of cross-linguistic influence have been identified to explain this variability. Typically, these have been discussed in relation to the *presence* of cross-linguistic influence, namely whether certain conditions have to be met for cross-linguistic influence to occur, and in relation to the *strength* of cross-linguistic influence, namely whether under certain circumstances the effect size of cross-linguistic influence increases.

In this study, we focus on four factors frequently studied in relation to cross-linguistic influence: (1) the type of surface overlap between bilingual children's languages, (2) the language domains involved, (3) language dominance, and (4) children's age. Whilst other factors, such as input quality (e.g., Paradis & Navarro, 2003) and economy principles (e.g., Gavarró, 2003; Serratrice et al., 2009), have also been argued to predict cross-linguistic influence, the number of studies investigating these variables is more limited and hence they are not included here. In the following four subsections, we discuss each of the factors of interest in more detail. We will end this section by discussing other reasons why there is such variation in results between and within studies on cross-linguistic influence.

Predictors of cross-linguistic influence

Surface overlap

One factor argued to predict the presence of cross-linguistic influence in bilingual children is the type of overlap between children's languages. According to Hulk and Müller (2000; Müller & Hulk, 2001), there has to be ambiguity in the child's language input for cross-linguistic influence to occur: if a certain structure in language A can be analysed (by the child) by either syntactic analysis X or Y and language B provides evidence for analysis X only, language B may reinforce the use of that analysis in language A, resulting in quantitative cross-linguistic influence. In other words, a certain type of overlap between children's languages is *necessary* for cross-linguistic influence to occur (see Döpke, 1998 for a similar proposal). Hulk and Müller's overlap hypothesis is usually referred to in terms of surface or structural overlap. Whilst some authors make an explicit distinction between the two terms (e.g., Nicoladis, 2006; Schmitz, Patuto, & Müller, 2012), most use them interchangeably to refer to the same construct. We use *surface overlap* throughout.

Hulk and Müller's overlap condition describes a situation of *partial overlap* (e.g., Unsworth, 2003). There is optionality in language A – due to ambiguity in the input – and in language B one of these options is the preferred option. As a consequence, cross-linguistic influence is predicted to go unidirectionally from language B to language A. For example, in Persian, compounds can either be left- or right-headed (e.g., ***bee**honey* for honeybee versus *head**ache***). In English, compounds can only be right-headed (e.g., Foroodi-Nejad & Paradis, 2009). As a consequence, in Persian-English bilingual children English may reinforce the use of right-headed compounds in Persian, leading to their overproduction. Following the surface overlap condition, situations of *complete overlap* (i.e., where bilingual children's two languages behave identically) and *no overlap* (i.e., where they behave completely differently) should not result in cross-linguistic influence, however.

Whilst some scholars have found cross-linguistic influence in the direction predicted by surface overlap (e.g., Austin, 2007; Haznedar, 2007; Hulk & Müller, 2000), others have not; or they have found evidence of cross-linguistic influence in the absence of surface overlap (e.g., Argyri & Sorace, 2007; Foroodi-Nejad & Paradis, 2009; Nicoladis & Gavrila, 2015).

Language domain

A second factor that has been argued to predict the presence of cross-linguistic influence in bilingual children is the language domain of the morphosyntactic property tested. Hulk and Müller proposed that, in addition to surface overlap, cross-linguistic influence only occurs in the domain where syntax interfaces with pragmatics, the so-called C-domain (e.g., Hulk & Müller, 2000; Müller & Hulk, 2001). An example is children's use of subject pronouns in a null subject language (e.g., Argyri & Sorace, 2007; Serratrice, 2007; Sorace et al., 2009). Null subject languages allow both overt and null pronouns in subject position. However, the choice of a pronoun depends on discourse-pragmatics principles (e.g., Carminati, 2002). In particular, whilst a null pronoun is typically used to refer back to the topic of the discourse, an overt pronoun signals a shift in discourse topic. Consequently, subject pronoun use in null subject languages has been argued to be at the interplay of syntax and (discourse-)pragmatics (e.g., Sorace & Serratrice, 2009; Sorace et al., 2009). However, Hulk and Müller did not rule out other domains at the interface with syntax (e.g., Hulk & Müller, 2000; Müller & Hulk, 2001, p. 2). Non-interface areas, such as purely syntactic language properties, were predicted to be unaffected (e.g., compounding; Foroodi-Nejad & Paradis, 2009; Nicoladis, 2002; root infinitives; Hulk & Müller, 2000).

Whilst some researchers have found evidence for Hulk and Müller's proposal (e.g., Argyri & Sorace, 2007; Haznedar, 2007), others have also found cross-linguistic influence in other domains, especially syntax-semantics (e.g., genericity and specificity; Serratrice et al., 2009; indefinite object scrambling; Meroni et al., 2017). Furthermore, cross-linguistic influence has not always been attested when discourse pragmatics were involved (e.g., Argyri & Sorace, 2007). Moreover, cross-linguistic influence in purely (morpho)syntactic properties of language is also attested (e.g., Argyri & Sorace, 2007; Nicoladis, 2012; Strik & Pérez-Leroux, 2011).

Language dominance

A third factor that has been related to cross-linguistic influence is language dominance. Bilingual children typically have a dominant and a weaker language (e.g., Grosjean, 1982). What counts as a child's dominant language can be defined in various ways, for example as the language a child is most proficient in (e.g., Unsworth, Chondrogianni, & Skarabela, 2018). Language dominance has been observed to predict both the presence and the strength of cross-linguistic influence. Some studies have found cross-linguistic influence to be unidirectional and, thus, to predict the direction of cross-linguistic influence, namely from children's dominant language into their non-dominant language (e.g., Argyri & Sorace, 2007; Yip & Matthews, 2000). Others have shown cross-linguistic influence to be bidirectional and to be present regardless of languages' dominance status. However, some studies found language dominance to predict the strength of cross-linguistic influence. To be more precise, the weaker the language was children have been tested in, the stronger the effect of cross-linguistic influence (e.g., Foroodi-Nejad & Paradis, 2009; Kidd, Chan, & Chiu, 2015; Nicoladis, 2006). At the same time, others have found no effects of language dominance (e.g., Foroodi-Nejad & Paradis, 2009; Nicoladis, 2002; Unsworth, 2012).

Age

A final factor observed to affect the presence and strength of cross-linguistic influence is age. Earlier studies of cross-linguistic influence were typically corpus studies with very young bilingual children (often before the age of four) investigating the development of a certain morphosyntactic property over a longer period of time (e.g., Döpke, 1998; Müller & Hulk, 2001; Serratrice et al., 2004). As already discussed, in those studies cross-linguistic influence was evident during time periods where bilingual children's acquisition was slower or faster than monolingual peers', and where bilingual children used qualitatively different structures than monolingual peers. Importantly, these studies suggested that cross-linguistic influence is a developmental phenomenon which, with sufficient language exposure, disappears over time (e.g., Döpke, 1998; Müller & Hulk, 2001; Paradis & Genesee, 1996).

In more recent experimental work, researchers have explored cross-linguistic influence in older bilingual children (e.g., Daskalaki, Chondrogianni, Blom, Argyri, & Paradis, 2019; Kaltsa et al., 2019). In an early study, Argyri and Sorace (2007) found evidence for cross-linguistic influence in seven-to-nine-year-old children, and others have found cross-linguistic influence to remain stable with age (e.g., Bosch & Unsworth, 2020; Nicoladis, 2002, 2003). This suggests that rather than being an exclusively developmental phenomenon, cross-linguistic influence may be part and parcel of being bilingual (e.g., Nicoladis, 2006, 2012; Serratrice, 2013, 2016). At the same time, some experimental studies have found the effect of cross-linguistic influence to diminish (e.g., Serratrice et al., 2009; Sorace et al., 2009; Unsworth, 2012) or even increase with age (e.g., Nicoladis & Gavrila, 2015). As a consequence, it is currently still unclear whether cross-linguistic influence is primarily a developmental phenomenon, mostly found in young bilingual children, or persists with age. Furthermore, as pointed out to us by an anonymous reviewer, age can be an index of language input and might therefore correlate with the (cumulative) amount of language exposure children receive in their two languages. We return to this latter point in the discussion.

In sum, despite or perhaps even because of the considerable body of experimental research on the topic, there is as yet no consensus about the circumstances under which cross-linguistic influence occurs. The presence of cross-linguistic influence and effects of its predictors vary across studies. In the next section, we discuss several explanations for this variability.

Accounting for variability across studies
First of all, study designs vary considerably in task set-up, morphosyntactic properties, and language pairs tested. Furthermore, the context of bilingual acquisition varies both within and across studies (e.g., in terms of input and age of onset). Whilst this variation across studies is necessary to detect whether there is a robust effect of cross-linguistic influence, study differences may influence the extent of cross-linguistic influence in unknown ways.

Second, surface overlap and language dominance have been defined and operationalized in many ways. With regard to surface overlap, some studies have based their predictions about surface overlap on the perspective of the adult language (e.g., Argyri & Sorace, 2007), whereas other studies focused on the (monolingual) child's point of view (e.g., Pirvulescu, Pérez-Leroux, Roberge, Strik, & Thomas, 2014). For example, whilst adult native speakers of English might not allow left-headed compounds (e.g., *beehoney* referring to the insect), monolingual children might consider such orders possible in English (e.g., Foroodi-Nejad & Paradis, 2009). The first scenario may have resulted in the underestimation of options available to the child and hence to the potentially incorrect classification of certain morphosyntactic properties as not overlapping between children's languages.

With regard to language dominance, authors have measured dominance differently, and operationalized it as both a categorical and continuous variable (e.g., Hervé, Serratrice, & Corley, 2016; Nicoladis, 2002; Unsworth, 2012). For example, some divided bilingual children into dominance groups (e.g., Argyri & Sorace, 2007; Foroodi-Nejad & Paradis, 2009), whereas others included a continuous measure of dominance, such as percentage of language exposure or scores on some measure of language proficiency in their analyses (e.g., Bosch & Unsworth, 2020; Nicoladis, 2002). These differences in definitions and operationalizations may explain why studies have found different effects of surface overlap and language dominance.

Third, the absence of a significant effect in situations where cross-linguistic influence has been predicted should not be interpreted as absence of cross-linguistic influence. Instead, non-significant effects are to be expected due to random error. If we assume that the power of studies investigating cross-linguistic influence is 80%, then there is a 20% chance that studies fail to detect a significant effect of cross-linguistic influence when it is in fact there. Scholars often interpret non-significant effects incorrectly as the absence of an effect (cf. Borenstein, Hedges, Higgins, & Rothstein, 2009; Brysbaert, 2019). Instead, what is essential is whether the direction of the non-significant effects was consistent with cross-linguistic influence. Given that it is common for studies on cross-linguistic influence to test relatively few bilingual children (e.g., Foroodi-Nejad & Paradis, 2009; Strik & Pérez-Leroux, 2011), many studies probably even had a lower power level than 80%. Underpowered studies and random variables could therefore also explain why some studies have failed to find significant effects of surface overlap, language domain, language dominance and age, whilst others have.

2.2 The present study

The aim of the present study is to conduct a meta-analysis that systematically assesses cross-linguistic influence and its predictors. Such a meta-analysis allows us to go beyond problematic differences between studies, because summary effect sizes are calculated for relevant variables by averaging across studies. In this way, effects of cross-linguistic influence can be investigated in much larger groups of children than in individual studies. Furthermore, a meta-analysis can provide information on whether variation in the effect of cross-linguistic influence between studies appears to be random (i.e., is due to random error), or systematic (i.e., relates to predictor variables; Borenstein et al., 2009). Finally, a meta-analysis allows us to statistically test the role of predictor variables (e.g., Borenstein et al., 2009; Hedges & Olkin, 1985).

In this study, we address the following research questions:

RQ1 To what extent is there cross-linguistic influence in bilingual children at the level of morphosyntax and how consistent is this effect across studies?

Given that cross-linguistic influence has been attested in various studies (e.g., Serratrice, 2013), we expect to find an average effect size of cross-linguistic influence that is significantly larger than zero. At the same time, we expect considerable variation across studies due to differences in experimental designs. Nevertheless, findings from studies should generally be consistent with cross-linguistic influence.

RQ2 To what extent does surface overlap affect the strength and presence of cross-linguistic influence?

We hypothesize that if the strength of cross-linguistic influence is affected by surface overlap, its effect will be stronger in situations of partial surface overlap – when one language can reinforce a partially overlapping morphosyntactic structure in the other language – compared to situations without surface overlap. If, however, surface overlap is a *necessary* condition for cross-linguistic influence to occur at all (e.g., Hulk & Müller, 2000), the effect of cross-linguistic influence will be significant only in situations with partial surface overlap and not in situations of no surface overlap.

RQ3 To what extent does language domain affect the strength and presence of cross-linguistic influence?

If language domain affects the strength of cross-linguistic influence, we expect cross-linguistic influence to be stronger for morphosyntactic properties that interact with discourse pragmatics compared to properties in other language domains. However, if the interaction between morphosyntax and discourse pragmatics is *necessary* for cross-linguistic influence to be present (e.g., Hulk & Müller, 2000), the effect of cross-linguistic influence will only be significant in this domain and not in others.

RQ4 To what extent does language dominance affect the strength and presence of cross-linguistic influence?

If language dominance affects the strength of cross-linguistic influence (e.g., Argyri & Sorace, 2007; Foroodi-Nejad & Paradis, 2009), cross-linguistic influence should be stronger from children’s dominant language into their non-dominant language than vice versa. If language dominance affects the presence of cross-linguistic influence, we hypothesize that cross-linguistic influence will be unidirectional from children’s dominant language into the non-dominant language (e.g., Yip & Matthews, 2000). Hence, the effect of cross-linguistic influence should only be significant in children’s non-dominant language and not in children’s dominant language.

In sum, for the role of surface overlap (RQ 2), language domain (RQ 3) and language dominance (RQ 4), we formulated both a weaker and a stronger version of our hypotheses. The weaker hypothesis considers the predictors’ effect on the strength of cross-linguistic influence. The stronger hypothesis considers its effect on the presence of cross-linguistic influence. We tested these hypotheses in two ways: (i) by using the authors’ categorization of surface overlap, language domain and language dominance; and (ii) by categorizing the predictors ourselves. This second way of coding had the advantage, first of all, that it allowed for systematicity in terms of the definition and operationalization of cross-linguistic influence across studies; and, second, effect sizes could be taken into account for predictors not explicitly tested by the authors themselves.

RQ5 How does cross-linguistic influence develop with age?

We hypothesize that if cross-linguistic influence is a developmental phenomenon (e.g., Hulk & Müller, 2000; Paradis & Genesee, 1996), the effect of cross-linguistic influence should become weaker as children grow older. This is in line with studies that have found cross-linguistic influence to become weaker or disappear with age (e.g., Serratrice et al., 2009; Sorace et al., 2009). In contrast, if cross-linguistic influence is part and parcel of being bilingual, no significant effect of age on the strength of cross-linguistic influence should occur (e.g., Bosch & Unsworth, 2020; Nicoladis, 2002; 2003).

2.3 Method

Literature searches

We began by building a systematic inventory of studies investigating cross-linguistic influence in bilingual children (see Figure 2.1). We selected studies that measured differences in bilingual and monolingual children's performance on a certain language task for specific morphosyntactic properties and interpreted their findings in relation to cross-linguistic influence. The following additional inclusion and exclusion criteria were applied:

Inclusion criteria

- Children were simultaneous and early sequential bilinguals, i.e., age of onset for both languages was before the age of 4;0 (e.g., Genesee, Paradis, & Crago, 2004; McLaughlin, 1978; Unsworth, 2013);
- Children were no older than 10;0 at the time of testing;
- The study presented original data.
- The study contained data from at least two bilingual and two monolingual children.

Exclusion criteria

- Studies with bimodal bilingual children, adoptees and children with a developmental language disorder;
- Priming and narrative studies.

We first searched Google Scholar for articles using various terms for cross-linguistic influence in combination with "bilingual children" (July, 2018; see Figure 2.1). We selected the first 980 returns for each term. In a second step, all articles were screened by two coders on the basis of titles and abstracts with respect to aforementioned criteria. Subsequent full-text screening revealed that the vast majority of articles were irrelevant for our purposes because they either focussed on bilingual adults or on a topic other than cross-linguistic influence. In cases of disagreement, a third person acted as arbiter. If necessary, we contacted the study's authors to check whether our criteria were met. In a third step, we searched the references cited in the selected articles for additional relevant studies, and we asked a number of experts in the field whether they knew of any studies not yet included.

In total, our search yielded 37 studies that met our inclusion criteria, and for 28 of these, we contacted authors for additional data (see below). In 15 cases, our request was met. For one study (Nicoladis, 2002), we were able to deduce the necessary information from reported statistics. For another study (Sorace et al., 2009), we estimated data from figures reported in the paper. For 11 studies, no sufficient data could be retrieved. Our final dataset therefore consisted of 26 studies.

Data coding

All but one of the 26 studies reported multiple comparisons between bilingual and monolingual children. For example, some studies investigated cross-linguistic influence in both bilingual children's languages or for various morphosyntactic properties. Furthermore, some studies explored the behaviour of various bilingual groups, split up, for example, by age, country of residence, language dominance profile, and age of first exposure (e.g., Argyri & Sorace, 2007; Meir, Walters, & Armon-Lotem, 2017; Serratrice et al., 2009; Serratrice, Sorace, Filiaci, & Baldo, 2012; Strik & Pérez-Leroux, 2011). We entered each comparison as a separate row in a spreadsheet, yielding 187 unique datapoints.[1]

---

[1] In some situations not all comparisons reported in the selected studies met our initial selection criteria, either because a bilingual group was added as control group for another bilingual group, rather than as a test case of cross-linguistic influence (i.e., the Spanish-Dutch bilingual group in Serratrice et al., 2009; 2012; Sorace et al., 2009) or because a specific condition was not at the level of morphosyntax (i.e. the stressed and unstressed pronouns in English in Serratrice et al., 2012). Datapoints belonging to such comparisons were excluded.

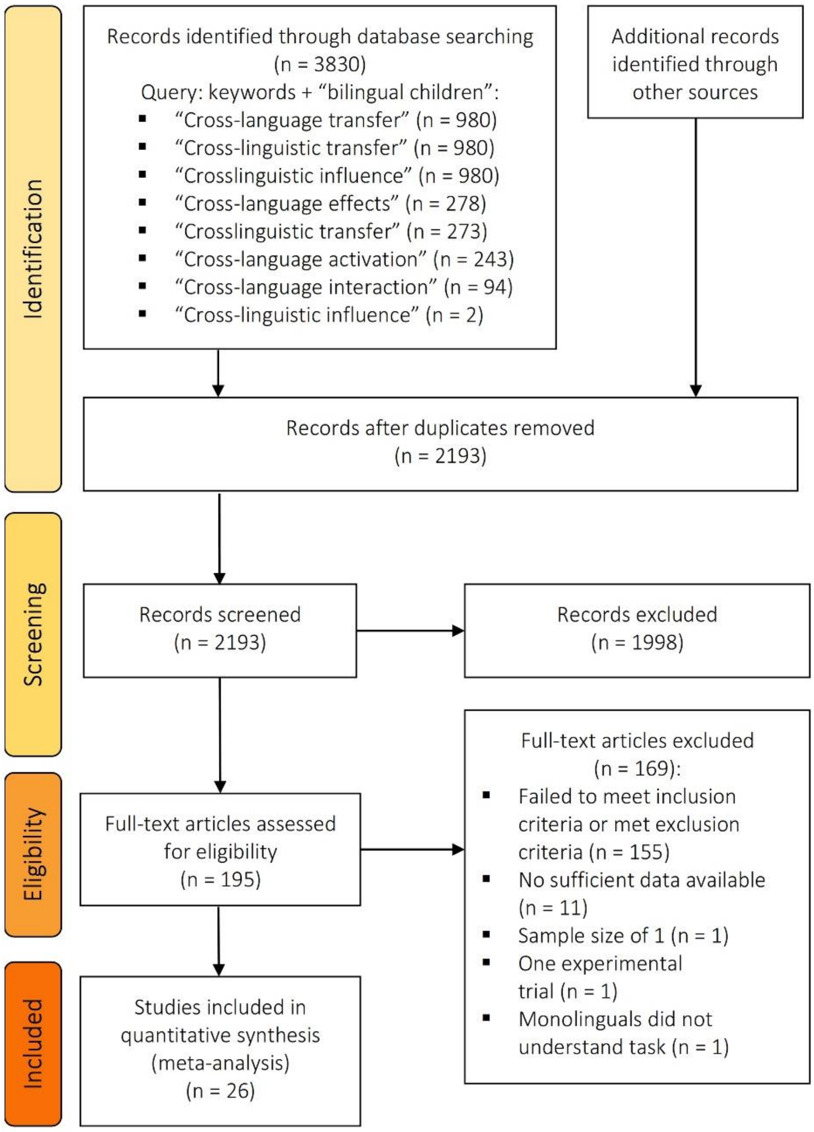


Figure 2.1. Flow chart showing selection process of experimental studies on cross-linguistic influence in simultaneous bilingual children following Preferred Reporting Items for Systematic Reviews and Meta-Analyses by Moher, Liberati, Tetzlaff, Altman, and Group, 2009.

Subsequently, we coded each datapoint for a number of characteristics, including task design, language tested and morphosyntactic property, adapting a template provided by Metalab (http://metalab.stanford.edu; e.g., Bergmann et al., 2018). We coded for our variables of interest, namely surface overlap, language domain, language dominance and age, and indicated whether a datapoint was considered as a testcase of cross-linguistic influence. The complete dataset is publicly available in the Data Archiving and Networked Services (DANS) repository (van Dijk, van Wonderen, Koutamanis, Kootstra & Unsworth, 2021).

Testcases of cross-linguistic influence

A datapoint was coded as a testcase of cross-linguistic influence in two steps. We coded first whether authors made explicit predictions about cross-linguistic influence (yielding 145 datapoints), and second, the direction of the predicted effect. For example, Foroodi-Nejad and Paradis (2009) elicited the production of compounds in Persian-English bilingual children. They predicted that *if* cross-linguistic influence were to take place, the children should use more right-headed compounds in Persian and/or more left-headed compounds in English compared to their monolingual peers. Hence, for the Persian task the direction of cross-linguistic influence predicted by the authors was coded as "more right-headed compounds" and for the English task as "more left-headed compounds". Cross-linguistic influence was predicted for a total of 103 datapoints. For 42 datapoints, authors predicted no effect of cross-linguistic influence. This was typically the case when bilingual children's languages patterned similarly for the morphosyntactic property under study (i.e., complete overlap). Hence, in those situations, bilingual children were predicted to behave similarly to monolingual children and datapoints were not included in the analyses.[2] Unfortunately, authors did not always formulate explicit predictions about cross-linguistic influence for each possible comparison (42 datapoints; e.g., Gathercole, Laporte, & Thomas, 2005; Sorace et al., 2009).

[2] Sometimes authors stated multiple – conflicting – hypotheses for the same datapoint (23 datapoints). For example, Serratrice and colleagues (2009) predicted unidirectional cross-linguistic influence from Italian to English based on Hulk and Müller's surface overlap condition (e.g., Hulk & Müller, 2000; Müller & Hulk, 2001). However, they also formulated an alternative hypothesis based on economy considerations, which predicted unidirectional cross-linguistic influence in the opposite direction: from English to Italian. In addition, some authors predicted cross-linguistic influence according to one theory and no cross-linguistic influence according to another theory (e.g., Serratrice et al., 2009). In all conflicting situations, we categorized datapoints as a testcase of cross-linguistic influence (in the direction(s) indicated by the authors).

To avoid inconsistencies across studies, we therefore applied a second, more neutral way of coding for testcases of cross-linguistic influence. We first identified every datapoint for which the authors had made no explicit predictions about cross-linguistic influence or for which they predicted no cross-linguistic influence (84 datapoints). We then checked whether the morphosyntactic property involved differed between bilingual children's languages. This was done based on information that was provided in the articles. If a morphosyntactic property was identical between bilingual children's languages, we predicted no cross-linguistic influence. These datapoints were then excluded from the analyses.[3] If a morphosyntactic property differed between bilingual children's languages, we coded the datapoint as a testcase of cross-linguistic influence. With regard to the direction of cross-linguistic influence for these newly identified datapoints, we predicted that bilingual children would use a certain morphosyntactic structure more than their monolingual peers if this structure was preferred in their other language. Our second way of coding yielded 40 possible testcases of cross-linguistic influence in addition to those datapoints for which the authors themselves predicted cross-linguistic influence. We now turn to how we coded our moderator variables.

Predictors of cross-linguistic influence

Surface overlap. Our first predictor of interest was operationalized in two ways: (i) the authors' definition of overlap when based on Hulk and Müller's (2000) overlap hypothesis; and (ii) our own definition of overlap. The first operationalization yielded 35 datapoints that were identified by the authors as a situation of surface overlap, 60 situations of no surface overlap, six

[3] Initially, we wanted to compare the average effect size of cases of complete overlap to testcases of cross-linguistic influence in order to shed light on the distinction between cross-linguistic influence and a more general effect of bilingualism. This turned out to be impossible, however, because the direction of individual effect sizes differs: for testcases of cross-linguistic influence the direction of Hedges' *g* can be positive (consistent with cross-linguistic influence) or negative (inconsistent with cross-linguistic influence), whereas for cases of complete overlap there is no such distinction. Consequently, effect sizes would either always be positive or negative for cases of complete overlap. As a result, we deemed a comparison between cases of complete overlap and testcases of cross-linguistic influence to be uninformative. For a similar reason, situations with *complete overlap* were not included in the surface overlap analyses. Even if bilingual children would be found to behave differently from their monolingual peers in *complete overlap* situations, the effect size will never be positive (indicating cross-linguistic influence).

situations of complete surface overlap, and one situation where the authors first identified the situation as surface overlap but later argued that their task may in fact have tested a situation of complete overlap instead. From these 102 datapoints, we excluded datapoints for which no predictions could be made about cross-linguistic influence, i.e., when the predicted direction of cross-linguistic influence could not be inferred (6 datapoints) and when there was complete overlap between languages (6 datapoints). This left us with a total of 90 datapoints.

This way of coding turned out to have two disadvantages, however. First, many authors either did not explicitly discuss their study in relation to Hulk and Müller's overlap hypothesis (65 datapoints) or made no explicit predictions (20 datapoints). Second, for those datapoints that could be included in the analysis, authors varied as to whether they defined surface overlap in terms of (i) the adult- or the child-language system (we will elaborate on this in the Discussion); and (ii) a narrowly defined morphosyntactic context versus a broader context (see A2.1 in the appendix for an explanation of narrow versus broad scope). To deal with these issues, we recoded all datapoints using the same criteria, namely based on the adult system and using a narrow scope. This not only allowed us to code for surface overlap in a uniform way, it also meant we could include datapoints from studies that made no explicit predictions about surface overlap. Datapoints were either coded as a situation of partial overlap (41 datapoints), a situation of no overlap (67 datapoints) or a situation of complete overlap (27 datapoints). For the remaining 52 datapoints, no unambiguous classification was possible. A detailed illustration of how datapoints were classified is provided in the appendix (A2.2).

Language domain. With respect to our second predictor of interest, language domain, we coded whether authors indicated which language domains were involved in the distribution of a certain morphosyntactic property, for example syntax and pragmatics, or syntax and semantics. This was mentioned explicitly for 70 datapoints only: 43 datapoints involved discourse pragmatics, 20 datapoints were identified as purely (morpho)syntactic, and 7 datapoints involved semantics and not discourse pragmatics.

In an attempt to include more datapoints, we tried to systematically code for language domains ourselves. This turned out to be problematic. Hulk and Müller's (2000, p. 228) original definition was "the interface between two modules of grammar, and more particularly at the interface between pragmatics and syntax". This definition is rather vague. Sorace and Serratrice (2009, p. 196) provide a more specific definition: "the distribution of the morphosyntactic construction of interest must be regulated by the interface with discourse pragmatics". This latter definition can be straightforwardly applied to cases such as the distribution of null and overt subjects in languages such as Greek and Italian (e.g., Argyri & Sorace, 2007; Sorace et al., 2009), but in many other cases it was almost impossible to determine when discourse pragmatics were *not* involved. Hence, we decided to only analyse those datapoints for which language domain was mentioned by the authors.

Language dominance. Our third variable of interest, language dominance, was coded in two ways: (i) depending on the definition of the authors; and (ii) depending on the societal language of the bilingual children tested. The first way of coding was as follows: if authors classified a group of bilingual children as dominant in one of their two languages, we classified them as dominant in the target language ("target language"; 24 datapoints) or in the non-target language ("other language"; 23 datapoints) depending on which language was tested. Children considered balanced by the authors were coded as "balanced" (14 datapoints), and in cases where authors wrote that bilingual children's dominance patterns varied, we coded dominance as "mixed" (6 datapoints). Information about dominance was not always provided. Consequently, language dominance was coded for a subset of datapoints only (67 in total).

Because language dominance was not consistently operationalized across studies, we also assessed children's language profile in a more systematic way by coding whether or not the target language was also the language of the society where the bilingual children lived. Although we realize that this is only a rough proxy of language dominance (Hervé et al., 2016; Unsworth et al., 2018), it does provide a more objective measure of children's language experience which could be coded for most studies.

Societal language was operationalized as the majority language of the country or area where the children were living and was coded as follows: if the societal language was the target language of a study, language dominance was coded as "target" (81 datapoints), if not, it was coded as "other" (97 datapoints). In one study (Hervé & Lawyer, unpublished manuscript) bilingual children came from different countries with different societal languages ("mixed"; 8 datapoints) and in one study (Nicoladis & Gavrila, 2015) there was no clear distinction in status for the children's two languages ("both", 1 datapoint).

Age. Our fourth predictor of interest, age, was coded as mean age in months. In all studies the bilingual and monolingual children had similar mean ages except for the older bilingual group tested by Strik and Pérez-Leroux (2011). The age range of this bilingual group (6;05–7;11) and its monolingual control group (4;07–5;08) did not overlap. Such a large difference could have been problematic for our moderator analysis because younger children are typically less accurate on a language task than older children. Therefore, effects of cross-linguistic influence may both be exaggerated or minimized, depending on whether cross-linguistic influence is predicted to result in facilitation or delay. To avoid these effects, we excluded the datapoints from the older group of bilingual children in Strik and Pérez-Leroux (2011) from our analysis of age (4 datapoints). In addition, we also excluded the results from the English task in Serratrice et al. (2009; 8 datapoints), because it was unclear which results belonged to the younger and older age group tested.

Effect sizes

Effect size estimates

We calculated the standardized effect size Hedges' *g*, and its variance, for each datapoint (e.g., Borenstein et al., 2009; Hedges, 1981; all calculations were taken from Lakens, 2013, version 4.2). Each effect size was based on the differential mean of a bilingual and a monolingual group on a certain measure. The larger the difference in means between groups and the smaller their standard deviations were, the larger Hedges' *g*. In addition, we calculated the variance of Hedges' *g*. This indicated the precision of an effect size (e.g., Borenstein et al., 2009). The larger the group sample sizes were, the smaller the variance and the more precise the corresponding effect size. In the meta-analysis, the more precise an effect size was, the more weight it was assigned.

The sign of the effect sizes indicated whether differences in scores found between bilingual and monolingual children were consistent with cross-linguistic influence. If the difference between a bilingual and a monolingual group was in the predicted direction the corresponding effect size was positive. If, on the other hand, there was a difference between a bilingual and monolingual group, but in the opposite direction than predicted (i.e., inconsistent with cross-linguistic influence), the corresponding effect size was negative. If bilingual and monolingual children had a similar score, the effect size was zero. We illustrate the interpretation of positive and negative effect sizes with two examples from Nicoladis (2006).

Nicoladis (2006) investigated cross-linguistic influence in adjective-noun orders in French-English bilingual children. In French, most adjectives typically appear postnominally (e.g., *une pomme vert*, "an apple green") whereas some typically appear prenominally (e.g., *une grande pomme*, "a big apple”). In English, adjectives should – with a few exceptions – appear in prenominal position only (e.g., *a green/big apple*). Hence, Nicoladis predicted cross-linguistic influence from English into French to result in more prenominal adjectives in bilingual children’s speech production compared to monolingual French peers. She elicited adjective-noun pairs in two conditions: (i) with typical French postnominal adjectives; and (ii) with typical French prenominal adjectives. She found bilingual children to produce the prenominal adjective order with postnominal adjectives in French more often than monolingual children. This difference between groups was consistent with cross-linguistic influence from English and therefore received a positive effect size. In addition, bilingual children also placed prenominal adjectives more often in postnominal position than French monolingual children. This observation was inconsistent with cross-linguistic influence from English. Consequently, the effect size received a negative sign.

Data dependency

Effect sizes in our dataset were often not independent because they belonged to similar studies, similar groups of children or similar morphosyntactic properties investigated. Following Fernández-Castilla et al. (2020, 2019) we controlled for dependencies in the data by a multiple level cross-classified random effects model. In this model, we added three random effects for observation (i.e., an individual datapoint), namely (i) a random intercept of observation nested in experimental task, which, in turn, was nested in data collection,[4] (ii) a random intercept of observation nested in group of bilingual children, which, in turn, was nested in data collection, and (iii) a random intercept of observation nested in morphosyntactic property. All models in the paper used this random-effects structure.

Our random effect structure accounted for most dependencies in our dataset. One exception concerned those datapoints for which outcomes of different groups of bilingual children were compared to the same outcome from a group of monolingual children. To simplify our dataset, we collapsed means and standard deviations for datapoints belonging to different groups of bilingual children and similar groups of monolingual children by calculating weighted means and pooled standard deviations (e.g., Hoyt & Del Re, 2018). This resulted in a total of 128 datapoints. In our analyses of language

[4] The same task in two different languages within the same data collection was coded as two separate tasks. In addition, we decided to nest participant groups and tasks in data collection rather than in study because data from Serratrice et al. (2009; 2012) and Sorace et al. (2009) were collected within the same data collection.

dominance, we used uncollapsed datapoints in those situations where separate bilingual groups had different dominance patterns. This yielded 176 datapoints.

Data analyses

All analyses were conducted using the *rma.mv* function from the *metafor*-package (version 2.4-0; Viechtbauer, 2010) in R (version 3.6.3; R Core Team, 2020). For all analyses, the aforementioned random effect structure was applied. We performed two types of analyses: general analyses of the weighted mean effect size, and predictor analyses. First, we tested the average weighted mean effect size of cross-linguistic influence twice: (i) for those datapoints for which the authors made explicit predictions about the direction of an effect of cross-linguistic influence, and (ii) for all datapoints which we identified as possible testcases of cross-linguistic influence.

Second, we conducted separate moderator analyses with surface overlap, language domain and language dominance as predictors to investigate their effect on the strength and presence of cross-linguistic influence. With respect to surface overlap, effect sizes were compared twice: (i) between situations of surface overlap and no surface overlap as defined by the authors of the studies based on Hulk and Müller's (2000) overlap hypothesis, and (ii) between situations of partial overlap and no overlap as defined by us (see footnote 3 for an explanation why we could not take into account situations of complete overlap). If the difference between either of these surface overlap situations was significant, we tested whether surface overlap affected the presence of cross-linguistic influence. This was done by assessing whether the effect of no overlap and partial overlap was significantly larger than zero.

With respect to language domain, we conducted one analysis in which we compared the effect size of cross-linguistic influence for morphosyntactic properties that interacted with discourse pragmatics to the effect size for morphosyntactic properties that did not interact with discourse pragmatics. If this difference was significant, we assessed whether the effect size of cross-linguistic influence in each situation was significantly larger than zero.

To test the effect of language dominance on the strength of cross-linguistic influence, effect sizes were compared twice: (i) between groups of children that were categorized as either dominant in the language tested or in the other language by the authors of the studies[5]; and (ii) between groups of children whose language of testing was the societal language and whose language of testing was not the societal language. If the difference between dominance categories was significant, we tested whether language dominance affected the presence of cross-linguistic influence. This was done by assessing whether the effect in the dominant and non-dominant language was significantly larger than zero.

Finally, with regard to age, a meta-regression was conducted with the mean age of the bilingual groups as continuous predictor of the effect size of cross-linguistic influence.

2.4 Results

Descriptive results

Our dataset consisted of 187 datapoints belonging to 750 unique bilingual children compared to 739 unique monolingual children. An overview of the characteristics of the studies in the dataset can be found in the appendix (A2.3). The majority of studies employed elicited production tasks. However, most observations in the dataset belonged to grammaticality judgement experiments. Only a few studies considered cross-linguistic influence in children's comprehension. There is considerable variation in the languages and linguistic properties tested. Although English has received most attention, there are many observations for other languages, too. Moreover, the language combinations under study were even more varied, with 17 unique language combinations. With regard to the linguistic properties tested, a large proportion investigated cross-linguistic influence in word order. Furthermore, quite a few studies focussed on null subjects and objects. However, the category with the most observations was genericity/specificity of plural noun phrases, even though only two studies tested for this property. Finally, with regard to the number of items tested per child, the majority of studies tested for cross-linguistic influence for a specific condition in less than 10 items. Eleven studies tested 6 items or less. Only six studies tested more than 20 items.

An overview of the characteristics of the bilingual groups in the dataset can be found in the appendix (A2.3) as well. The most frequently tested age group for bilingual children was on average four years old. Only

[5] Datapoints belonging to children whose dominance profile was described as mixed or balanced were not included, due to low numbers of datapoints (mixed: 6; balanced: 14).

two studies considered cross-linguistic influence in three-year-olds. With regard to the number of children studied, it is noteworthy that 17 studies compared groups of bilingual children to monolingual peers with a sample size of less than 20 for the bilinguals, and, in seven studies, with a sample size of less than 10. Although the majority of studies tested groups of 20 or more bilingual children, the majority of observations in our dataset belong to smaller sample sizes.

Cross-linguistic influence: average effect size and consistency

Figure 2.2 shows the datapoints per study for which either the authors or we predicted cross-linguistic influence (see the supplementary material on the LOT publications webshop for forest plots with information about the morphosyntactic property and the language combination tested split out by task type).[6] The majority of effect sizes were larger than zero (73 datapoints), consistent with cross-linguistic influence. However, there were also a number of negative effect sizes (24 datapoints), which was inconsistent with cross-linguistic influence. Furthermore, the effect size of cross-linguistic influence varied between and within studies.

In our first analysis of the average effect size of cross-linguistic influence, we included only the 79 datapoints for which the authors of the studies explicitly predicted cross-linguistic influence. Effect sizes ranged from -1.24 to 2.66. The random effects model revealed a significant small to medium average effect size of $g = 0.46$ ((0.22, 0.71), $p < .001$).

In our second analysis, we included an additional 34 effect sizes (a total of 113) previously identified as possible testcases of cross-linguistic influence. Now, the effect sizes ranged from -1.37 to 2.66. The random effects model revealed a significant small to medium average effect size of $g = 0.39$ ((0.21, 0.56), $p < .001$), slightly smaller than the average effect size in the first analysis.[7]

[6] The distribution of the subset of effect sizes for which the authors explicitly predicted cross-linguistic influence was very similar to the distribution of effect sizes in Figure 2.2. Therefore, we decided to present the full set only.

[7] An anonymous reviewer was concerned that the average weighted effect size was not entirely reliable because we collapsed effect sizes of different task types. We did test for the effect for task type (elicited production, judgements and comprehension) in a moderator analysis, but this did not yield a significant effect. Outcomes of subset analyses for each task type can be found in the appendix (A2.4).

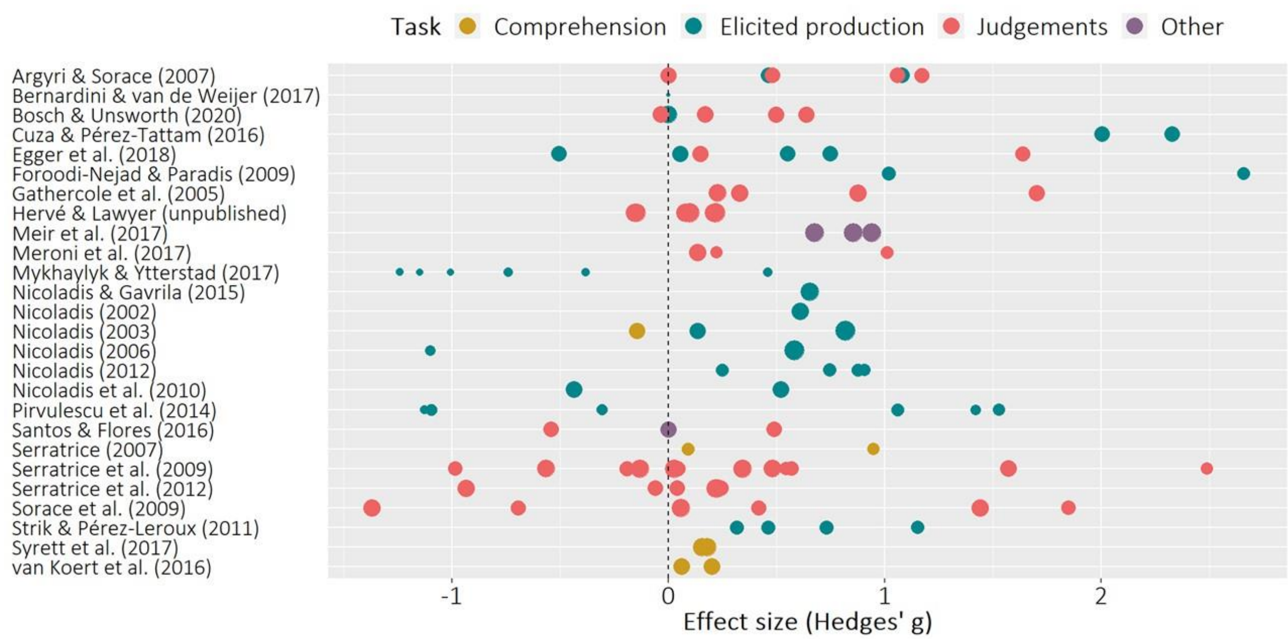


Figure 2.2. Effect sizes per study belonging to datapoints that were identified as possible testcases of cross-linguistic influence. The dot sizes indicate the precision of each effect size. Positive effect sizes reflect differences between bilingual and monolingual groups consistent with cross-linguistic influence. Negative effect sizes reflect differences inconsistent with cross-linguistic influence.

We further investigated the distribution of effect sizes in the second analysis using a funnel plot (Figure 2.3). In this plot, datapoints are plotted with their effect size on the x-axis and their standard error on the y-axis. The vertical line represents the average effect size. Datapoints with a smaller standard error are predicted to be scattered closer to the average effect size than datapoints with a greater standard error, as indicated by the diagonal lines. If studies with significant results are more likely to be published than studies with null results (publication bias), this should be reflected in an asymmetrical distribution of datapoints in the funnel plot: there should be more datapoints at the bottom right side of the distribution than at the bottom left side (e.g., Rothstein, Sutton, & Borenstein, 2005). We do not see this distribution in Figure 2.3. Instead, there seemed to be some asymmetry in the opposite direction, namely there were a number of effect sizes at the lower left side of the distribution.

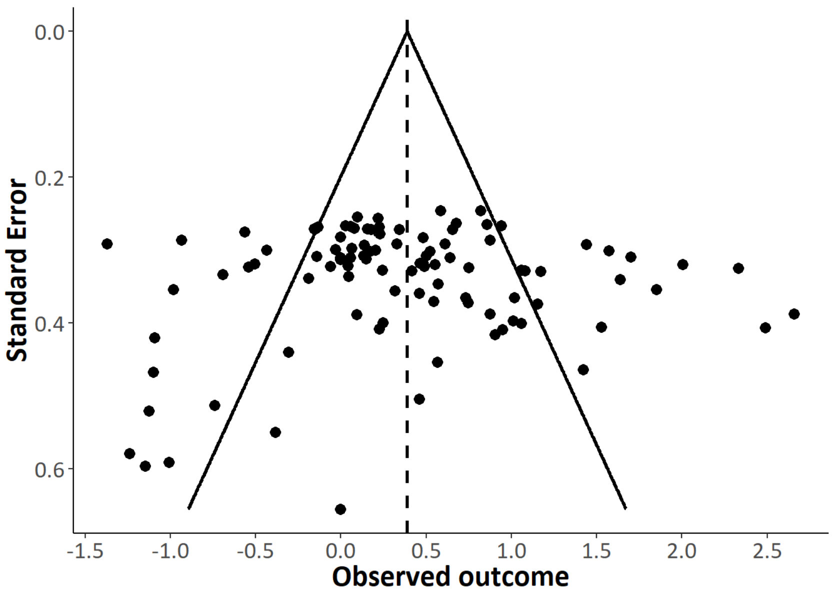


Figure 2.3. Funnel plot with observed effect sizes plotted on the x-axis and their standard errors on the y-axis.

Figure 2.3 also revealed quite some horizontal scatter of datapoints, a signal of heterogeneity in the data (e.g., Sterne et al., 2011). This was confirmed by the significant test of heterogeneity of the model ($Q(112) = 505.00$, $p < .001$), which indicated that part of the variance in the data could not be explained by random error alone. This means that there must be other factors at play that account for differences in effect sizes. We tested whether this variance could be explained by our predictors of interest.

Analyses of predictors of cross-linguistic influence

We analysed the effect of our predictors by means of meta-regressions (e.g., Viechtbauer, 2010). All predictor analyses were conducted with positive effect sizes only. Negative effect sizes reflected divergent behaviour between bilingual and monolingual children that was inconsistent with cross-linguistic influence. We will discuss possible reasons for negative effects sizes in the Discussion. Regardless of what causes negative effect sizes in our dataset, interpreting them is difficult, and their presence might muddy our predictor analyses. Therefore, we decided to leave out negative effect sizes from further analyses. Moderator tests were conducted separately for our predictors of interest.

Surface overlap

The first analysis took into account those datapoints for which the authors made predictions about the presence or absence of cross-linguistic influence based on Hulk and Müller's (2000) overlap hypothesis. Overall, the average effect size for surface overlap situations was slightly larger (M = 0.69, SD = 0.81, *range* = 0-2.66, n = 20) than the average effect size of situations without surface overlap (M = 0.54, SD = 0.58, *range* = 0-2.49, n = 31). However, this difference was not significant as shown by the moderator test of surface overlap (Q_M (1) = 1.78, p = .182).

The second analysis compared the average effect size of those datapoints that we identified as partial overlap situations versus no overlap situations. The average effect size for partial overlap was slightly larger (M = 0.76, SD = 0.72, *range* = 0-2.66, n = 17) than the average effect size of no overlap (M = 0.62, SD = 0.64, *range* = 0-2.49, n = 42). However, the difference in effect size between partial overlap and no overlap situations did not reach significance either (Q_M (1) = 0.37, p = .541).

Language domain

The average effect size of morphosyntactic properties at the domain of discourse pragmatics (M = 0.30, SD = 0.38, *range* = 0-1.17, n = 19) was slightly smaller than the average effect size of morphosyntactic properties at other domains (M = 0.39, SD = 0.40, *range* = 0-1.06, n = 17). This difference was not significant, however (Q_M (1) = 0.05, p = .832).

Language dominance

In the first analysis, we compared effect sizes between children that were tested in their dominant language against children that were tested in their non-dominant language, as defined by the authors. Effect sizes were larger when children were tested in their non-dominant language (M = 0.53, SD = 0.90, *range* = 0-3.42, n = 21) compared to their dominant language (M = 0.35, SD = 0.52, *range* = 0-1.65, n = 23), Q_M (1) = 4.35, p = .037. However, when inspecting Cook's distance and DFBETA values one datapoint was identified that had a relatively large effect on the outcome of the model (g = 3.42, standardized residual, z = 3.00). We therefore re-ran the model without this datapoint. Effect sizes were still slightly larger when children were tested in their non-dominant language (M = 0.39, SD = 0.62, *range* = 0-1.80, n = 20) compared to their dominant language (M = 0.35, SD = 0.52, *range* = 0-1.65, n = 23). However, this difference no longer reached significance (Q_M (1) = 2.05, p = .152). This showed that the initial significant effect was carried by the effect size that was removed.

In the second analysis, the effect of societal language was tested. The average effect size of cross-linguistic influence was larger in those situations where the language of testing was not the societal language (M = 0.82, SD = 1.31, *range* = 0-7.54, n = 61) compared to when it was the societal language (M = 0.49, SD = 0.51, *range* = 0-2.05, n = 57), Q_M (1) = 6.86, p = .009. When inspecting Cook's distance and DFBETA values, two influential effect sizes were identified (g = 7.54, standardized residual, z = 6.83; and g = 5.16, standardized residual, z = 4.82). Without these two effect sizes, the difference in effect sizes between children tested in their non-societal language (M = 0.64, SD = 0.80, *range* = 0-3.64, n = 59) and in their societal languages (M = 0.49, SD = 0.51, *range* = 0-2.05, n = 57) was not significant but the trend was in the same direction (Q_M (1) = 3.36, p = .067). Furthermore, the estimated effect size of children tested in their non-societal and in their societal language was significantly larger than zero (non-societal language: B = 0.70, SE = 0.12, (0.47-0.93), p < .001; societal language: B = 0.52, SE = 0.12, (0.29-0.75), p < .001), indicating that the effect size of cross-linguistic influence was significant in the direction of the societal language into the non-societal language and in the direction of the non-societal language into the societal language.

Age

Figure 2.4 presents the distribution of the effect sizes by the average age of the bilingual groups by task type (107 datapoints). Two observations can be made. First, studies with younger children (< 6;0) in our dataset typically employed elicited production tasks to test for cross-linguistic influence. In older children, on the other hand, cross-linguistic influence was more often measured through judgement tasks. Second, the older children were, the smaller the effect of cross-linguistic influence became. This pattern was not significant, however (Q_M (1) = 0.46, B = -0.003, SE = 0.004, p = .497).

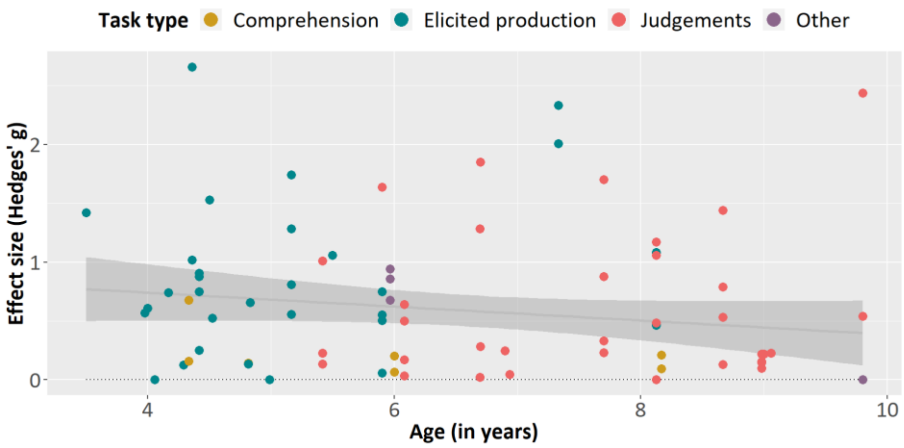


Figure 2.4. Effect sizes as a function of children's age (in years) by task.

2.5 Discussion

In this study, we systematically reviewed previous research on cross-linguistic influence in bilingual children by means of a meta-analysis. Our aim was to assess the strength of cross-linguistic influence by generalizing over differences in methodology and linguistic properties. In addition, we investigated the effect of previously identified predictors of cross-linguistic influence, namely surface overlap, language domain, language dominance, and age. A total of 26 studies met our inclusion criteria, which resulted in a total of 187 datapoints. Subsets of the available datapoints were included in the analyses testing our predictors of interest. In this section we first discuss our findings, before using them to make a number of recommendations for future studies on cross-linguistic influence.

Cross-linguistic influence: average effect size and data consistency

We assessed the presence, strength and consistency of cross-linguistic influence in previous research with bilingual children. We hypothesized that (i) there would be an overall significant effect of cross-linguistic influence, and (ii) the effect sizes of individual studies would be consistent with cross-linguistic influence. Our findings fully supported our first hypothesis and partially supported our second hypothesis.

A significant summary effect of cross-linguistic influence was observed across studies. Bilingual children's languages influence each other at the level of morphosyntax, in line with the general consensus in the literature (e.g., Serratrice, 2013). Our analyses revealed a small to moderate effect size, as reflected in a Hedges' *g* between 0.39 and 0.45. The moderate but not strong effect size indicates that although bilingual children's languages can influence each other, they generally behave in language-specific ways similar to monolingual children (e.g., Nicoladis, 2002; Paradis & Genesee, 1996). This effect size may serve as a benchmark for future studies on cross-linguistic influence, and stimulate researchers to conduct power analyses for determining the necessary minimum sample size (e.g., Cohen, 1988).

We observed that authors did not always formulate comprehensive predictions about cross-linguistic influence. Instead, some studies focussed on certain conditions only, even when more were tested. Possibly, authors might have felt inclined to solely report significant or large effects. Indeed, the summary effects of cross-linguistic influence was slightly larger when we only took those datapoints into account for which authors made explicit predictions. There was no evidence for a publication bias in our funnel plot, however. Alternatively, authors might have focussed on conditions that offered clearest support for their theoretical perspective on cross-linguistic influence. Regardless of the reason, incomplete predictions made studies less transparent and outcomes more difficult to interpret and compare to outcomes of other studies.

Finally, most but not all datapoints in our dataset were consistent with cross-linguistic influence. Out of 113 effect sizes, 73 showed a difference between bilingual and monolingual children consistent with cross-linguistic influence. Thus, given the variety of study designs in our dataset, cross-linguistic influence can present itself regardless the type of task set-up used or the linguistic property and language combination tested. However, 24 effect sizes went in the opposite direction and the magnitude of cross-linguistic influence varied largely across and within studies. We address this in the next sections.

Predictors of cross-linguistic influence

Surface overlap

We hypothesized that cross-linguistic influence should be stronger in situations of surface overlap versus no surface overlap. If surface overlap is a necessary condition for cross-linguistic influence, the average effect size should be significant only in situations of surface overlap. (e.g., Hulk & Müller, 2000; Müller & Hulk, 2001). This turned out not to be the case, neither when surface overlap was coded based on authors' definitions, nor when systematically coded by us based on the adult system. The average effect size of cross-linguistic influence was not significantly different in situations of surface overlap and situations of no surface overlap.

Our analyses show that surface overlap as presently defined does not significantly affect the size of the cross-linguistic effect. However, on the basis of our results it would be inappropriate to conclude that effects of cross-linguistic influence are unaffected by *any* type of surface overlap. It is possible that when surface overlap is defined in terms of ambiguity and optionality in the child's developing system, cross-linguistic influence may still be found (e.g., Hulk & Müller, 2000; Müller & Hulk, 2001).

Take, for example, Foroodi-Nejad and Paradis' (2009) study. Their results can either be interpreted as evidence for *or* against the surface overlap hypothesis, depending on how surface overlap is defined. If surface overlap is based on the adult system, English constitutes a situation of no surface overlap with Persian, because English only allows right-headed compounds (whereas Persian allows both left- and right-headed compounds). If surface overlap is based on the child system, however, English might actually constitute a situation of surface overlap with Persian, because English monolingual children have been found to sometimes produce ungrammatical left-headed compounds (e.g., Foroodi-Nejad & Paradis, 2009; Nicoladis, 2002). Foroodi-Nejad and Paradis (2009) found that Persian-English children produced more left-headed compounds in English than monolingual peers. On a definition of surface overlap based on the adult system, this means that there was cross-linguistic influence in a situation of no overlap. However, on a definition based on the child system, these results constitute cross-linguistic influence in a situation of surface overlap.

Because we and most authors of the studies in our dataset defined surface overlap based on the adult system, the number of situations of surface overlap in the meta-analysis might have been underestimated. Unfortunately, we were unable to test for effects of surface overlap based on the child system as most studies provided too little information to do so. Further systematic investigation of the role of surface overlap when defined in terms of child versus the adult language system is needed.

Language domain

We hypothesized that when morphosyntax interacts with discourse pragmatics, the size of cross-linguistic influence should be stronger than in other domains. If cross-linguistic influence is only present in a domain with such an interaction (e.g., Hulk & Müller, 2000; Müller & Hulk, 2001), the average effect size of cross-linguistic influence should be significant only in this domain. This hypothesis was not borne out: there was no significant difference in effect sizes for morphosyntactic properties whose distribution was governed by discourse pragmatics compared to other morphosyntactic properties. These findings suggest that cross-linguistic influence can occur irrespective of language domain (*contra* Hulk & Müller, 2000; Müller & Hulk, 2001).

However, it proved difficult to categorise morphosyntactic properties into specific domains, as there was often no clear line between situations in which discourse pragmatics are and are not involved (e.g., Montrul, 2011; Sorace, 2011). An alternative proposal would be to focus on computational complexity (e.g., Hopp, 2009; Sorace, 2011). Under such an account, certain morphosyntactic properties should be more sensitive to cross-linguistic influence due to their relative complexity (along the lines of Hulk and Müller's original proposal), and cross-linguistic influence could occur regardless of the language domain involved. Indeed, several studies have found evidence for the involvement of computational complexity in cross-linguistic influence (e.g., Gavarró, 2003; Strik & Pérez-Leroux, 2011).

In sum, rather than linguistic domain, computational complexity may be a more relevant predictor of cross-linguistic influence. Further investigation on this topic is needed to test this idea systematically.

Language dominance

With respect to language dominance, we hypothesized that if language dominance affects the size of cross-linguistic influence the average effect size of cross-linguistic influence would be larger from the dominant into the non-dominant language rather than the other way round. If cross-linguistic influence is from the dominant into the non-dominant language only, we predicted the effect of cross-linguistic influence to be significant in that situation only. Two analyses were conducted. We first analysed those datapoints for which the authors categorized the bilingual group as either dominant or non-dominant in the language tested. Subsequently, we operationalized language dominance in terms of the societal language. Evidence was found for the first, but not the second part of the hypothesis.

Cross-linguistic influence was stronger from children's societal language into the non-societal language than vice versa. Furthermore, the effect of cross-linguistic influence from children's non-societal language into their societal language was significantly larger than zero. In contrast, when the authors' dominance groups were analysed, no evidence for an effect of language dominance was found. Taken together, these results suggest that language dominance, as operationalized by societal language, does not predict the *presence* of cross-linguistic influence, but rather its *strength*.

The absence of an effect of dominance in the first analysis is most likely due to the differences in how authors categorized children in dominance groups. Typically, three measurements were used to assess children's dominance profile: amount of language exposure (and use), lexical proficiency, and fluency ratings by parents or teachers. Some studies combined (some of) these measurements when categorizing children into dominance groups (e.g., Foroodi-Nejad & Paradis, 2009; Pirvulescu et al., 2014). Other studies used only one of these measurements (e.g., Argyri & Sorace, 2007; Serratrice, 2007). Different measures may lead to children being assigned to different dominance groups, however (Unsworth et al., 2018).

In sum, variation *within* dominance groups may have masked differences *between* dominance groups in the first analysis, resulting in the absence of a significant effect of language dominance. Future studies should therefore consider testing for the effect of dominance on cross-linguistic influence by exploring different proxies for language dominance separately.

Age

With regard to age, two hypotheses were formulated: (i) if cross-linguistic influence is a developmental phenomenon, the average effect size of cross-linguistic influence should become smaller over age; (ii) if, on the other hand, cross-linguistic influence is part and parcel of being bilingual, the average effect size of cross-linguistic influence should not differ with age.

Our results were consistent with the second hypothesis. The average effect size of cross-linguistic influence did not significantly change over age. This is in line with those previous studies that found cross-linguistic influence to remain present in older bilingual children (e.g., Argyri & Sorace, 2007; Bosch & Unsworth, 2020; Kaltsa et al., 2019).

Our findings are in contrast with spontaneous production studies with very young children that attested cross-linguistic influence only during a certain phase in language development (e.g., Döpke, 1998; Hulk & Müller, 2000). This could be explained by the different modalities tested with younger and older children. In our dataset, cross-linguistic influence in older groups of bilingual children was mainly tested by judgement tasks. Possibly, these studies detected subtle effects of cross-linguistic influence that were only present in older bilingual children's judgements of sentences and not in their (spontaneous) speech production. If this is correct, cross-linguistic influence may be less strong in older bilingual children's speech production than their judgements, but this needs empirical confirmation (cf. Argyri & Sorace, 2007; Kaltsa et al., 2019). It is also possible that some instances of cross-linguistic influence may be developmental in nature, whereas others are more persistent.

Two words of caution are required here. As pointed out to us by two anonymous reviewers, the effect of age on cross-linguistic influence might be more complex than it appears in the present study. First, bilingual children's age might serve as a proxy for relative exposure and as such for their language dominance. In particular, children might experience a switch in dominance from the home language to the societal language after starting school (e.g., Polinsky & Kagan, 2007). Consequently, the expected direction of cross-linguistic influence may change as children become older. Second, the relation between age and cross-linguistic influence may be modulated by the age of acquisition of the specific morpho-syntactic phenomenon in question. If cross-linguistic influence only occurs whilst children are in the process of acquiring the language property in question, then it is predicted to persist for properties that are acquired late (e.g., pronoun interpretation in languages like Italian and Greek; Papadopoulou, Peristeri, Plemenou, Marinis, & Tsimpli, 2015), whereas it should be less apparent for properties that are acquired early (e.g., Verb Second in Dutch and German; Wijnen & Verrips, 1998). When the same property is acquired at different rates in different languages (e.g., gender in Greek versus gender in Dutch; Egger, Hulk, & Tsimpli, 2018), this may lead to asymmetric effects of cross-linguistic influence in bilingual children acquiring those languages. By combining different morphosyntactic properties from different languages, we were unfortunately unable to disentangle effects of age from effects of age of acquisition. We encourage researchers to use the information in our dataset to conduct more in-depth analyses of age effects whilst at the same time pointing out that establishing the age of acquisition for each property in all of the relevant languages is by no means trivial. Our initial attempts to do so revealed that the necessary information was often unavailable or inconclusive.

Unexplained variation

Although some of the variance in effect sizes of cross-linguistic influence in our dataset could be explained by children's societal language, much of the variance remains unexplained, as does the observation that there were negative effect sizes. We deal with each of these issues in turn.

With respect to unexplained variance, a number of causes can be considered. First, part of the unexplained variance in effect sizes may be due to the operationalization of surface overlap and language dominance. If it had been possible to define those two constructs in a different, better way – as explained above – they might have had accounted for (more) variation in the data. Our observations that the average effect size of cross-linguistic influence in situations of partial overlap and in children's dominant language was slightly but not significantly larger than in situations of no overlap and in children's non-dominant language offer support for this view.

Second, part of the unexplained variance could potentially be attributed to different types of bilingual acquisition, as pointed out to us by an anonymous reviewer. In particular, whilst some of the children in our dataset were acquiring their languages in a *one parent, one language* situation, others were in families where both parents spoke the minority language at home. The context in which children acquire their languages is relevant for the (cumulative) amount of input children receive (e.g., Unsworth, Brouwer, de Bree, & Verhagen, 2019), which in turn, is related to their patterns of language dominance (e.g., Unsworth et al., 2018). Consequently, average effect sizes of cross-linguistic influence might differ depending on the type of bilingual acquisition involved. Although studies in our dataset often reported at least some information about the languages spoken at home, they did not always provide the (enough) relevant details. We therefore could not include the role of acquisition type in our analyses.

Third, more general effects of bilingualism could contribute to differences in performance between bilingual and monolingual children. For example, bilingual children might have performed less accurately on certain tasks compared to monolingual peers because of comparatively reduced input in their two languages or because they experienced increased processing demands having to deal with two languages instead of one (e.g., Pirvulescu et al., 2014; Sorace & Serratrice, 2009). While this latter claim remains a moot point, it is possible that general effects of bilingualism may have had a greater impact on certain morphosyntactic properties than others, and especially on those properties that require a large amount of input to be acquired or that are difficult to process. This could, in part, explain why effect sizes differed across studies. In other words, effect sizes in our dataset may not have been pure reflections of cross-linguistic influence, but may have consisted of other effects as well.

Some evidence for general bilingualism effects in bilingual children comes from a study by Sorace and colleagues (2009). They tested bilingual and monolingual children's choices of null and overt subject pronouns in Italian. They included a group of Spanish-Italian bilingual children. Spanish and Italian are both null subject languages and have similar preferences regarding subject pronoun choices (e.g., Sorace et al., 2009; but cf. Filiaci, 2010). Regardless of the overlap between languages, Sorace and colleagues found Spanish-Italian children to be less accurate in their pronoun choices than their monolingual Italian peers. Consequently, they argued that more general bilingualism effects, such as processing difficulties, affected children's pronoun choices, rather than cross-linguistic influence (also see Sorace & Serratrice, 2009 for an extensive discussion).

Fourth, all effect sizes in our dataset came from elicited production studies and offline judgement and comprehension tasks. More recent accounts of cross-linguistic influence have suggested that cross-linguistic influence is the result of language co-activation during sentence processing (e.g., Bosch & Unsworth, 2020; Nicoladis, 2006, 2012; Serratrice, 2013, 2016). As the strength of language co-activation may have varied from study to study – for example, due to differences in children's language experiences – cross-linguistic influence may not always have surfaced in children's production and offline judgements and comprehension.

Special attention should be paid to the presence of negative effect sizes. These effect sizes represented differences between bilingual and monolingual children inconsistent with cross-linguistic influence. For example, we predicted that *if* cross-linguistic influence was to affect French-English bilingual children's placement of prenominal adjectives in French in Nicoladis (2006), bilingual children should be more accurate in their production of adjective-noun strings than monolingual peers. This is because English only allows for prenominal adjectives. However, bilingual children (age-matched to the monolingual children) in Nicoladis (2006) placed prenominal adjectives in French in postnominal position almost 50% of the time, versus about 10% in the French monolingual group (g = -1.10, s = 0.22). Although it could be argued that this difference between groups was a coincidence, it seems unlikely to find such a large difference between groups if cross-linguistic influence were actually present.

To account for negative effect sizes, two explanations should be considered. First, cross-linguistic influence might sometimes have resulted in a different strategy than predicted by the authors or by us. It is typically expected that cross-linguistic influence reinforces the use of a morphosyntactic structure in one of the children's languages when it is preferred in their other language. An alternative account would be that bilingual children may sometimes try to differentiate between the morphosyntax of their languages by making their languages as different as possible (Döpke, 1998). In other words, bilingual children might adhere to canonical morpho-syntactic structures as much as possible to differentiate between languages. In the case of French, postnominal adjectives are more frequent than prenominal ones (e.g., Nicoladis, 2006). Perhaps some bilingual children in Nicoladis (2006) placed prenominal adjectives in French in postnominal position so frequently in order to contrast postnominal adjective-noun strings in French to prenominal adjective-noun strings in English. On this account, cross-linguistic influence may have led (some) bilingual children to behave in more language-specific ways than monolingual children.

It is also possible that general effects of bilingualism might explain negative effect sizes. For example, in some experiments bilingual children might have performed less accurately on a task compared to monolingual peers as a result of less input in the language tested (e.g., Pirvulescu et al., 2014). This could explain why the bilingual children in Nicoladis (2006) more often incorrectly placed prenominal adjectives in French in postnominal position than monolingual children, that is, they may not have heard enough input in French to establish the prenominal position as a consistent option in that language. If a bilingualism effect were indeed responsible for children's inconsistent behaviour with regard to cross-linguistic influence, the challenge for future studies would then be to disentangle those effects from effects of cross-linguistic influence, especially when predictions go in the same direction.

Recommendations

Facilitating reproducibility and cross-study comparisons

First of all, we recommend studies to formulate clear and testable hypotheses for each condition tested. Ideally, to make studies testing for cross-linguistic influence as transparent as possible and less vulnerable to bias, authors should take the following steps: (i) state for all conditions tested how children's languages are different or similar; and (ii) based on this first step, state for each condition *if* cross-linguistic influence could manifest itself, and, importantly, *what* this cross-linguistic influence should look like when there is cross-linguistic influence and when not. Furthermore, in order to make direct comparisons across studies possible, studies should report effect sizes.

Operationalising surface overlap and language dominance

Surface overlap and language dominance should be defined and operationalized in uniform and transparent ways. With regard to surface overlap, we recommend authors to take each of the following steps: (i) describe the morphosyntactic property under study in the adult system of bilingual children's languages, at both the level of the specific context tested as well as at a more general level (for example, subjects in Greek *wh*-embedded interrogatives are always postverbal (specific context) but in other contexts they can appear preverbally as well (general context)); (ii) describe how the morphosyntactic property is acquired by monolingual and, if the relevant information is available, by bilingual children, and describe whether there is optionality during acquisition; and (iii) formulate hypotheses regarding surface overlap and indicate whether these are based on optionality in the adult language or the child language (ideally both).

With regard to language dominance, the field should strive for a standard, uniform way to define dominance. As long as this is not available, we would recommend authors to measure and report effects of amount of language exposure/use, proficiency and societal language on cross-linguistic influence separately, for example, using existing questionnaires (e.g., ALDeQ – Paradis, Emmerzael, & Sorenson Duncan, 2010; BiLEC – Unsworth, 2013; PaBiQ – Tuller, 2015). This way, effects of these separate proxies of language dominance can be compared and better understood.

Cross-linguistic influence versus general effects of bilingualism

We recommend that studies differentiate effects of cross-linguistic influence from possible effects of bilingualism. For most studies in our dataset, it was impossible to determine whether effect sizes consistent with cross-linguistic influence were (partially) driven by more general effects of bilingualism as well (cf. Pirvulescu et al., 2014; Serratrice et al., 2009; 2012; Sorace et al., 2009). We therefore propose that future studies include, where possible, an appropriate bilingual control group (e.g., Kaltsa et al., 2019; Serratrice et al., 2009, 2012; Sorace et al., 2009). Crucially, without this bilingual control group, it may be impossible to determine whether differences between a bilingual and monolingual group should be attributed to cross-linguistic influence or to a more general bilingualism effect (for similar discussion concerning adult second language learners, see Jarvis, 2000).

We do realize that for practical reasons it is not always possible to add a control group. In these situations, we recommend authors to consider the introduction of multiple within-experiment conditions that test the same cross-linguistic effects in different ways, and/or the inclusion of matched control-conditions in which only general bilingual effects would be expected (e.g., complete-overlap conditions).

Effect sizes from these studies could then be used to calculate a more precise average effect size of cross-linguistic influence.

Sample size and power

Ideally, future studies should consider the minimum sample size of children necessary to obtain a significant effect of cross-linguistic influence. If the true effect size of cross-linguistic influence is 0.39, then, a sample size of at least 82 children would be necessary in the bilingual and monolingual control group to detect this effect (for an alpha level of .05 and a beta level of .80). If the true effect size is 0.45 a minimum sample size of 62 children per group would be necessary (calculations were performed with G*power; Faul, Erdfelder, Lang, & Buchner, 2007). This means that with just one exception (Meir et al., 2017), all the studies in our dataset will likely have been underpowered. In fact, the vast majority of studies did not even test half of the participants required. We do realize that increasing sample sizes is easier said than done, especially given the relative scarcity of certain bilingual populations and the labour intensity of the data collection process. One solution to the power problem would be for researchers to collaborate when possible (Brysbaert, 2019).

Apart from testing more participants, researchers could aim to increase the sensitivity of their studies by decreasing error variance as much as possible. For example, by keeping background variables, such as age, proficiency, and amount of exposure – or, if not possible, type of acquisition - as constant as possible between bilingual children and by increasing the numbers of items tested (Brysbaert, 2019; see also Quené (2010) for a further discussion how to increase the sensitivity of a study). For example, in 19 studies in our dataset at least some of the reported group means were based on less than 10 items per condition and in four studies there were even less than five items. This might have resulted in less precise outcomes – and therefore decreased power – compared to studies with more items per condition. Furthermore, many studies in our dataset reported rather broad language proficiency and/or exposures ranges for bilingual children (e.g., Cuza & Pérez-Tattam, 2016; Foroodi-Nejad & Paradis, 2009; Serratrice et al., 2009). It is possible that children with very different language profiles show different effects of cross-linguistic influence from other children, especially given our finding that language dominance affects the strength of cross-linguistic influence. Combining results from children with very different backgrounds might therefore increase the noise in the data, decreasing the likelihood of differences between bilingual and monolingual scores reaching significance. Moreover, one solution frequently adopted by authors of splitting children into different groups decreases the sample size, again resulting in a loss of power. As an alternative, authors could strive to select bilingual children with as similar linguistic background as possible to obtain more precise group effects.

Finally, our estimation of a minimum sample size of 62 to 82 children per group is based on the average effect size of cross-linguistic influence from studies for which it is unclear to what extent a more general effect of bilingualism was at play. Other factors might have affected the effect size of cross-linguistic influence that we were unable to test for in this meta-analysis and hence the effect size reported here may be an over- or underestimation. In the latter case, smaller minimum sample sizes would be required for a properly powered study. Future studies following our recommendations are necessary to clarify this issue further.

Understudied areas of cross-linguistic influence

Finally, we recommend conducting additional studies on cross-linguistic influence in children's comprehension. The majority of studies in our dataset were concerned with elicit production or judgement tasks and only a few studies concerned comprehension (Nicoladis, 2003; Serratrice, 2007; Syrett et al., 2017; van Koert, Koeneman, Hulk, & Weerman, 2016). It is therefore unclear whether the average effect sizes attested in our meta-analysis apply to cross-linguistic influence in comprehension as well.

Furthermore, all studies in our dataset focussed on cross-linguistic influence in children's production, offline judgements and comprehension. Until now, virtually no studies have focused on cross-linguistic influence during real-time morphosyntactic processing (cf. Hervé & Lawyer, unpublished manuscript; Lemmerth & Hopp, 2019). This, too, might have resulted in an underestimation of cross-linguistic influence attested in bilingual children. More online data are necessary to explore more subtle effects of cross-linguistic influence.

2.6 Conclusions

This meta-analysis is the first study to systematically assess the effect size of cross-linguistic influence in bilingual children and effects of surface overlap, language domain, language dominance and age. Overall, there was a significant effect of cross-linguistic influence across studies and its average effect size was small to moderate. Furthermore, the results of most of the studies were consistent with cross-linguistic influence. Cross-linguistic influence was stronger from children’s societal language into their non-societal language than vice versa. No effects were found for surface overlap – either as defined by the authors of the studies or based on the adult language system only – language domain, language dominance as operationalized by the authors of the studies, or age. These findings suggest that cross-linguistic influence is part and parcel of being bilingual and can manifest itself in various linguistic contexts. At the same time, our meta-analysis also shows that more systematic and standardized studies of cross-linguistic influence are necessary to fully understand this aspect of bilingual language development and use. This especially holds for the formulation of hypotheses about cross-linguistic influence and the operationalization of surface overlap and language dominance. We hope that the recommendations given here will serve as an impetus for the field to move towards a more standardized and unified way of testing for cross-linguistic influence and its predictors.

CHAPTER 3

Cross-linguistic influence in simultaneous bilingual children's online processing of long passives and Verb Second

A self-paced listening study

Abstract

To assess the presence of cross-linguistic influence, this study compared the processing of Dutch sentences by English-Dutch and German-Dutch bilingual and Dutch monolingual children in a self-paced listening task. We combined insights from studies on child bilingualism and adult second language acquisition. Sentence structures showing partial overlap between languages were investigated (long passives), as well as structures with complete or no overlap (verb second and verb third sentences). We found evidence for syntactic co-activation of overlapping structures in the form of inhibition during listening. Lexical and syntactic overlap between languages, and language dominance modulated effects. In particular, online cross-linguistic influence was visible only in the German-Dutch group. Furthermore, effects were most pronounced when structures partially overlapped and were absent in non-overlapping structures. Effects of online cross-linguistic influence became stronger the more German-dominant children were. Our results indicate that syntactic co-activation across languages affects sentence processing in bilingual children.

Based on: van Dijk, C.N., Dijkstra, T., & Unsworth, S. (under review). Cross-linguistic influence during online sentence processing in bilingual children.

3.1 Introduction

An intensively studied topic in child bilingualism is the influence of a (morpho)syntactic property in bilingual children's one language on their other language. As an example, consider a French-English bilingual child saying *apple green* instead of *green apple*, where French word order (Noun-Adjective) is used in English (Nicoladis, 2006). Studies on such cross-linguistic influence have traditionally investigated the interdependency of bilingual children's syntactic systems (e.g., Paradis & Genesee, 1996) and have focused on children's production, offline judgements and comprehension of sentences (see Serratrice, 2013, for an overview). Various factors have been argued to affect the presence and strength of cross-linguistic influence, including language overlap (e.g., Döpke, 1998; Hulk & Müller, 2000; Müller & Hulk, 2001) and language dominance (e.g., Argyri & Sorace, 2007; Yip & Matthews, 2000).

To date, however, hardly anything is known about cross-linguistic influence in bilingual children during real-time language processing (cf. Lemmerth & Hopp, 2019). In contrast, recent studies with bilingual adults have typically employed online techniques to study syntactic interactions between languages. These have shown subtle effects of cross-linguistic influence during sentence processing (e.g., Hopp, 2017; Runnqvist, Gollan, Costa, & Ferreira, 2013). For example, Runnqvist and colleagues (2013) found that Mandarin-English bilingual adults were faster to produce English possessive clauses that overlapped in word order between Mandarin and English (e.g., the woman's cat) than possessive clauses that did not (e.g., the cat of the woman). It is unclear whether similar effects are also present in bilingual children.

This study aimed to integrate insights from offline child and online adult studies to assess cross-linguistic influence during sentence processing in bilingual children. In doing so, we examined cross-linguistic influence during the processing of different Dutch word orders in English-Dutch and German-Dutch bilingual children by means of a self-paced listening task. Language overlap and dominance, previously identified predictors of cross-linguistic effects in offline comprehension and production experiments, were systematically manipulated to test their role online.

Cross-linguistic influence in bilingual children

Whilst bilingual children are mostly found to use their languages like monolinguals (e.g., Paradis & Genesee, 1996), cross-linguistic influence has been attested for various morphosyntactic properties, such as word order (e.g., Argyri & Sorace, 2007; Döpke, 1998), compounding (e.g., Foroodi-Nejad

& Paradis, 2009; Nicoladis, 2002), and pronoun use (e.g., Haznedar, 2007; Serratrice, Sorace, & Paoli, 2004). To account for the observed presence and absence of cross-linguistic influence, researchers have identified various predictors. Two are relevant to our study: language overlap and language dominance.

Language overlap and language dominance

With respect to language overlap, Hulk and Müller (2000; Müller & Hulk, 2001) proposed that for cross-linguistic influence to occur there should be overlap between the child's two languages. At the same time one of these languages should be ambiguous, in the sense that Language *a* allows for one structural analysis (X) and Language *α* for two (X and Y). In such a situation, Language *a* may reinforce the use of structural analysis X in Language *α* (Hulk & Müller, 2000; Müller & Hulk, 2001; also see Döpke, 1998). This can then result in either delay or acceleration during acquisition. In other words, if one language allows for one option only and another language for two options – at least from the perspective of the child – the language with one option might reinforce that option in the language with two options. Following Unsworth (2003), we refer to this condition as *partial overlap*. Findings with respect to the effects of language overlap are mixed. Some studies have found that cross-linguistic influence was predicted by *partial overlap* (e.g., Foroodi-Nejad & Paradis, 2009; Haznedar, 2007), while others did not (e.g., Argyri & Sorace, 2007). In addition, there are also studies that have observed cross-linguistic influence in the absence of overlap (*no overlap*; see Serratrice, 2013, for discussion). We are not aware of any offline comprehension or production studies investigating cross-linguistic influence in situations where a morphosyntactic property is completely shared between languages (*complete overlap*). We will discuss examples of *partial, no* and *complete overlap* in more detail when we describe our structures of interest.

Another factor found to predict cross-linguistic influence is language dominance (e.g., Argyri & Sorace, 2007). Language dominance has been operationalized in various ways, including (relative) language exposure and proficiency (e.g., Silva-Corvalán & Treffers-Daller, 2015). Here the observation is that language dominance can affect both the occurrence as well as the strength of cross-linguistic influence. Specifically, in some studies cross-linguistic influence has only been attested in Language *α* in children that were dominant in Language *a* (e.g., Serratrice et al., 2009; Yip & Matthews, 2000). In other studies cross-linguistic influence in Language *α* was stronger the more dominant children's Language *a* was compared to Language *α* (e.g., Bosch & Unsworth, 2020; Foroodi-Nejad & Paradis, 2009). There are, however, studies where no such relation has been observed (e.g., Nicoladis, 2002; Serratrice, Sorace, Filiaci, & Baldo, 2012).

Cross-linguistic influence during sentence processing

Language co-activation and priming

Data about cross-linguistic influence and its predictors in bilingual children have almost exclusively been collected using offline comprehension and production tasks. Nevertheless, it is especially findings from *online* tasks in adult second language (L2) acquisition that have inspired recent theories on cross-linguistic influence in bilingual children. In adult bilinguals, there is ample evidence for language non-selective lexical access (e.g., Dijkstra & van Heuven, 2002; Marian & Spivey, 2003). Even in a completely monolingual situation, the lexicons of both languages are activated during bilingual processing, resulting, for instance, in facilitatory or inhibitory effects of language overlap in 'special' items like cognates and false friends (e.g., Dijkstra, van Jaarsveld, & ten Brinke, 1998). Similarly, evidence for language co-activation at the sentence level has been observed in adult L2 learners, that is, in priming studies adults were more likely to use a certain structure in their one language after having heard this structure in their other language (e.g., Hartsuiker, Pickering, & Veltkamp, 2004; Loebell & Bock, 2003). These findings suggest that structures in adult L2 learners' one language can activate similar structures in their other language, even when the latter is not in use. Over time, this cross-linguistic priming has been argued to even result in shared syntactic structures between languages (Hartsuiker & Bernolet, 2017).

Various researchers have proposed that lexical co-activation and cross-linguistic syntactic priming are the mechanisms by which cross-linguistic influence take place in bilingual children. Some have argued that lexical co-activation during sentence processing in a bilingual child's one language activates syntactic structures in both languages (e.g., Nicoladis, 2006, 2012; Nicoladis, Rose, & Foursha-Stevenson, 2010). As a consequence, structures from both languages compete for selection. This competition can sometimes surface as cross-linguistic influence offline in the acceptance of a nontarget-like structure (e.g., Bosch & Unsworth, 2020). Others have suggested that language co-activation and priming can result in shared syntactic structures, as in adult second language learners (e.g., Serratrice, 2013, 2016). In line with these proposals, cross-language lexical co-activation has been observed in bilingual children (e.g., Von Holzen & Mani, 2012). There is also evidence of cross-language syntactic priming in children (e.g., Hsin, Legendre, & Omaki, 2013; Vasilyeva et al., 2010). However, so far there is no direct evidence for syntactic co-activation in bilingual children *during* sentence processing.

There are only a handful of studies that have investigated sentence processing in bilingual children. Most involve early second language learners rather than simultaneous bilingual children (cf. Lemmerth & Hopp, 2019), who have been the focus of the cross-linguistic influence literature. These studies tested whether children's online behaviour was comparable to that of monolingual peers when processing morphosyntactic properties in one of their languages that were either similar to those in their other language (e.g., Lemmerth & Hopp, 2019; Marinis, 2007; Marinis & Saddy, 2013), completely different from their other language (e.g., Chondrogianni, Vasić, Marinis, & Blom, 2015), or absent altogether (e.g., Chondrogianni, Marinis, Edwards, & Blom, 2015). Whilst most studies reported qualitatively similar results for monolingual and bilingual children none of these studies actually set out to test for cross-linguistic influence, with the exception of Lemmerth and Hopp (2019). As a consequence, specific factors known from production and offline comprehension studies to be relevant for cross-linguistic influence, such as language overlap and language dominance, were not included. The role of these factors in bilingual children's online processing therefore remains unclear.

Online cross-linguistic influence in adult bilinguals

To better understand how cross-linguistic influence can manifest itself during sentence processing and how language overlap and dominance are involved we now turn to online studies with adult bilinguals. With regard to language overlap, our aim is to see whether type of overlap affects online cross-linguistic influence, and if so, what the mechanism is behind this. We will first discuss evidence for online cross-linguistic influence in situations of *partial overlap*. Then we will turn to *no overlap* and *complete overlap*. Finally, we will discuss the role of language dominance.

In language production, online cross-linguistic influence has been found to facilitate language processing in a situation of *partial overlap* (Runnqvist et al., 2013). Runnqvist and colleagues (2103) investigated the timing of the production of English possessives by Mandarin-English, Spanish-English and monolingual English speakers. English allows for prenominal and postnominal possessives (i.e., *the man's stroller is pink* versus *the stroller of the man is pink*). Mandarin only allows prenominal possessives, whereas Spanish only allows postnominals. In other words, for the Mandarin-English bilinguals there is *partial overlap* between their two languages for prenominal possessives, and for the Spanish-English bilinguals for postnominal possessives. The authors found evidence for online cross-linguistic influence in the Mandarin-English group. Mandarin-English bilinguals were faster to produce partially overlapping prenominal structures in English than the Spanish-English bilinguals.

The authors discuss their finding in terms of frequency effects and make a direct connection between *partial overlap* – albeit not labelled as such – and cross-language priming. Given that the prenominal option is the only option in Mandarin, its relative frequency is higher compared to English, where a second option is available. The same holds for Spanish: the postnominal option is the only option in that language and hence is comparatively more frequent than the postnominal form in English. Runnqvist et al. hypothesized that if overlapping structures are connected or even shared between languages, the higher frequency of occurrence of the two possessive structures in Mandarin and Spanish should be inherited by English, to some extent at least. This, they argue, explains why the Mandarin-English bilinguals showed facilitation in processing the partially overlapping prenominal structure in English: the higher frequency of prenominal possessives in Mandarin boosted the activation of the same structure in English. The authors explained the absence of a facilitation effect in the postnominal structure in the Spanish-English group in frequency terms as well. They argued that because the postnominal structure in English is the dispreferred structure (and hence less frequent), priming over time from Spanish was not sufficient for visible facilitation effects during production.

Effects of partial overlap have also been investigated in online comprehension studies – albeit once again not discussed in these terms. In contrast to Runnqvist et al.'s (2013) production study, *partial overlap* in sentence processing seems to result in *less* efficient processing in L2 learners. For example, Foucart and Frenck-Mestre (2012) investigated sensitivity to gender agreement violations in adjective-noun phrases in French L2 learners with English as first language (L1) in a situation of *partial overlap* and *no overlap*. The canonical position of French adjectives is postnominal (e.g., *les chaises vertes*, "the chairs green"). However, some adjectives appear in prenominal position (e.g., *les petites chaises*, "the small chairs"). Furthermore, French adjectives have to agree in gender with the noun they modify (e.g., *les chaises vertes*, "the chairs$_{\text{FEM}}$ green$_{\text{FEM}}$" versus **les chaises verts*, "the chairs$_{\text{FEM}}$ green$_{\text{MASC}}$"). In contrast, adjectives in English are always prenominal (e.g., *the green chairs*) and gender agreement is absent. Hence, the word orders of adjective-noun combinations in French and English constitute a situation of *partial overlap* with the less frequent option in French (prenominal) overlapping with the only option in English.

Foucart and Frenck-Mestre measured participants' ERPs while they listened to gender violations in prenominal and postnominal adjective-noun pairs in French. Results showed that in the non-overlapping postnominal word order, L2 learners were as sensitive to gender violations as native speakers. However, in the overlapping prenominal order, L2 learners showed less sensitivity – as evidenced by the absence of a P600 effect. These results suggest that processing an L2 is affected by an L1 in a situation of *partial overlap* (also see, e.g., Alemán Bañón, Fiorentino, & Gabriele, 2014; but cf. Alemán Bañón et al., 2018).

An explanation of why *partial overlap* can result in less efficient processing in an L2 comes from Hopp (2017). In an eye-tracking study, English L2 learners with German as L1 were asked to read reduced relative clauses in English (e.g., *When the doctor Sarah ignored tried to leave*). The word order of such sentences corresponds to the canonical OV order in German (*Als die Ärtztin Sarah ignorierte*, "when the doctor ignored Sarah"). Hence, reduced relative clauses in English overlap in form with SOV clauses in German but not in meaning. Furthermore, the frequency of word orders differs between languages. SOV is the canonical and therefore a highly frequent word order in German subordinate clauses. The linear order of reduced relative clauses in English, in contrast, does not correspond to the canonical SVO structure of English (e.g., Lehmann, 1978). Hence, there is *partial overlap* in reduced relative clauses between English and German in the sense that the surface word order is similar in both languages, but structural representations differ in meaning and frequency. Note that our definition of *partial overlap* deviates here from the traditional definition used in the child bilingualism literature (e.g., Hulk & Müller, 2000).

Hopp found that German L2 learners of English slowed down when reading reduced relative clauses. He explained his findings in terms of syntactic co-activation and inhibition. He argued that the English word order of reduced relative clauses activated the canonical SOV order in German. As a consequence, processing resources had to be allocated to inhibit the German structure which was visible as a slowdown effect during reading.

A similar explanation can account for the findings by Foucart and Frenck-Mestre (2012) for prenominal adjectives in French. In English, the canonical position of adjectives is prenominal. In French, on the other hand, adjectives most frequently appear in postnominal position. Therefore, we expect that the less frequent prenominal adjective clauses in French in Foucart and Frenck-Mestre's experiment strongly co-activated the more frequent prenominal adjective-noun structure in English. As a consequence, French L2 learners had to allocate processing resources to inhibit co-activation of English. This, in turn, left fewer processing resources available to detect gender violations, resulting in different processing patterns in L2 learners compared to native speakers.

Evidence of cross-linguistic influence in adult bilingual sentence processing in situations of *no overlap* and *complete overlap* is less clear. On the one hand, studies like Foucart and Frenck-Mestre's (2012) suggest that L2 learners can process language similarly to native speakers in situations of *no overlap*. The same has also been found for situations of *complete overlap* (e.g., Alemán Bañón et al., 2014). At the same time, however, results from other ERP studies suggest that in both *no* and *complete overlap* situations language processing by L2 learners might be less automatized compared to native speakers (e.g., Andersson, Sayehli, & Gullberg, 2019; Gillon Dowens, Guo, Guo, Barber, & Carreiras, 2011). However, it is unclear whether differences between L2 learners and native speakers are due to properties of L2 learners' L1 or due to other factors such as proficiency and processing demands.

With regard to language dominance, studies on sentence processing in bilingual adults show mixed effects. On the one hand, findings suggest that cross-linguistic influence from the L1 in the L2 is strongest in less proficient L2 speakers (e.g., Alemán Bañón et al., 2018; Hopp, 2017). For example, inhibition effects from the L1 in Hopp (2017) were most pronounced in participants who had low L2 proficiency scores compared to other participants. Hence, these findings suggest that the more dominant adults are in the language not in use, the stronger online cross-linguistic influence becomes. On the other hand, online cross-linguistic influence has been attested in highly proficient L2 learners as well (e.g., Foucart & Frenck-Mestre, 2012; Gillon Dowens et al., 2011). In this vein, note that the cross-linguistic influence found by Runnqvist and colleagues (2013) was from the nondominant into the dominant language.

In sum, if we look at the adult bilingualism literature through the lens of the child bilingualism literature, we observe a number of parallels. First of all, the morphosyntactic properties of one language can influence the morphosyntactic properties of another language. This has been found in production and offline comprehension experiments with bilingual children and online experiments with bilingual adults, where it has been shown to result in less efficient processing. Second, language overlap and language dominance, predictors of offline cross-linguistic influence in bilingual children, seem to play a role in online cross-linguistic influence in bilingual adults as well. Specifically, online effects of cross-linguistic influence seem most pronounced in situations of *partial overlap* and in language learners who are dominant in the language not in use.

3.2 The present study

The goal of this study was to investigate the presence of cross-linguistic influence during sentence processing in bilingual children as dependent on language overlap and language dominance. We tested English-Dutch and German-Dutch bilingual children. Presenting the children with a self-paced listening task, we systematically manipulated the word order of long passives and verbs second structures in Dutch. Children processed sentences that either completely overlapped, partially overlapped or did not overlap between their two languages.

The long passive in Dutch, English, and German

Dutch passives can be formed with the auxiliary *worden* ("to become") and a past participle. The agent is expressed in an optional by-phrase following the past participle (V-PP; 1).

(1)	De	jongen	wordt	geduwd	door	het	meisje.
	the	boy	is being	pushed	by	the	girl

The canonical surface structure of Dutch declarative main clauses is SVO (e.g., Koster, 1975; Zwart, 2011). However, the order of prepositional phrases relative to the main verb is rather flexible (e.g., Koster, 1974). Consequently, the by-phrase can also precede (PP-V) the past participle (2):

(2)	De	jongen	wordt	door	het	meisje	geduwd.
	the	boy	is being	by	the	girl	pushed

"The boy is being pushed by the girl."

Both word orders are grammatical, but there is evidence that the V-PP word order is preferred by native-speaker adults (Bernolet, Hartsuiker, & Pickering, 2009).

Like in Dutch, the English long passive is formed by an auxiliary plus participle and a by-phrase:

(3) The boy is being pushed by the girl.

English has a rather rigid SVO word order (e.g., Lehmann, 1978), implying that in long passives the by-phrase directly follows the main verb. The equivalent of the Dutch PP-V word order is ungrammatical:

(4) *The boy is being by the girl pushed.

German passives are similar to Dutch and can be formed with the auxiliary *werden* (“to become”) and a past participle (e.g., Verhagen, 1992; 5).

(5)	Der	Junge	wird	von	dem	Mädchen	geschubst.
	the	boy	is being	by	the	girl	pushed

“The boy is being pushed by the girl.”

In contrast to Dutch, German is characterized by the so-called *Satzklammer* structure. In sentences with composite verb forms, the auxiliary and main verb form “sentence brackets” and NPs and PPs must appear between these brackets (e.g., Dürscheid, 2012). Consequently, in long passives the by-phrase precedes the verb (i.e., PP-V). Note that, although the V-PP word order is strictly speaking ruled out, movement (*Ausklämmerung*) to the right of the verb does sometimes occur (e.g., for stylistic reasons), especially in spoken language (e.g., Betz, 2008; Dürscheid, 2012; Haider, 2010; 6).

(6)	?Der	Junge	wird	geschubst	von	dem	Mädchen.
	the	boy	is being	pushed	by	the	girl

In line with Hulk and Müller (2000), the long passive constitutes a situation of *partial overlap* both between Dutch and English and between Dutch and German. Dutch has two possible structures (PP-V and V-PP) whereas English (V-PP) and German (PP-V) only have one. As a consequence, frequency distributions of the PP-V and V-PP word orders differ between languages. In Dutch, input of long passives children receive is divided between the PP-V and V-PP structure – with the V-PP structure being potentially the more frequent one. In German, the PP-V structure is the only grammatical option, just as the V-PP structure is in English. Hence, the PP-V and V-PP orders are by definition relatively more frequent in, respectively, German and English, than in Dutch.

The (long) passive is acquired relatively late in life (e.g., Bartke, 2004; Bever, 1970; Fox & Grodzinsky, 1998; Hirsch & Wexler, 2006; Verrips, 1996). A large-scale study on eleven different languages showed that Dutch-speaking, English-speaking, and German-speaking 5 year olds performed above chance, but not yet at ceiling level in their interpretation of long passives (Armon-Lotem et al., 2016; also see Hirsch & Wexler, 2006). We do not know of any data on the acquisition of the two word orders for long passives (PP-V and V-PP) in Dutch-speaking children.

Verb placement in Dutch, English and German

Our complete/no overlap property is verb placement in non-subject-initial sentences: in this case, Dutch and German overlap completely, and English and Dutch not at all.

Dutch is a V2 language. In main clauses, the second constituent position is occupied by the finite verb (e.g., Koster, 1975; Zwart, 2011). When a constituent other than the subject occurs in the first position, the verb raises past the subject and moves to the complementizer position, creating an XVSO word order (as in (7); e.g., Koster, 1975; Zwart, 2011). Sentences with XSVO order (henceforth V3, (8)) are ungrammatical in Dutch.

(7)	Gisteren	at	het	meisje	een	appel.
	yesterday	ate	the	girl	an	apple
	"Yesterday the girl ate an apple."					

(8)	*Gisteren	het	meisje	at	een	appel.
	yesterday	the	girl	ate	an	apple

As noted above, English declarative clauses maintain a strict SVO order, irrespective of the constituent in first position ((9) cf. (10); Lehman, 1978). Only under limited circumstances, for example, in wh-questions (11), does the subject-verb inversion occur and does the finite verb move to the second constituent position (Radford, 2004).

(9) Yesterday the girl ate an apple
(10) *Yesterday ate the girl an apple.
(11) When did the girl eat an apple?

Similar to Dutch, German is a V2 language and subject-verb inversion is required when a constituent different from the subject occupies first position (e.g., Haider, 2010):

(12)	*Gestern	das	Mädchen	aβ	einen	Apfel.
	yesterday	the	girl	ate	an	apple

(13)	Gestern	aβ	das	Mädchen	einen	Apfel.
	yesterday	ate	the	girl	an	apple

The V2 structure constitutes a situation of *complete overlap* between Dutch and German as both languages require V2 in main clauses. As a consequence, V2 should have similar frequency distributions in the two languages. V3 is ungrammatical in both languages, constituting a situation of *complete overlap* as well. Between Dutch and English, however, the V2 and V3 structures constitute a situation of *no overlap* as main clauses in English follow a XSVO order (i.e., V3), which is ungrammatical in Dutch.

In monolingual acquisition, Dutch- and German-speaking children hardly ever make errors when it comes to finiteness and verbal placement (e.g., Blom, 2003; Clahsen & Penke, 1992; Poeppel & Wexler, 1993; Wijnen & Verrips, 1998). This suggests that V2 is acquired relatively early. Errors in bilingual acquisition research provide evidence for cross-linguistic influence in children acquiring a non-V2 language alongside a V2 language. For example, V3 orders have been found in German-English simultaneous bilingual children (Döpke, 1998) and English-Dutch sequential bilingual children (Unsworth, 2016). In contrast, the simultaneous bilingual children in Bosch and Unsworth (2020) never produced V2 structures in their English. They did, however, accept V2 orders in English more often than monolingual controls. Taken together, these findings suggest that Dutch V2 and English V3 structures can influence each other in bilingual language development.

Research questions and hypotheses

Our first research question is to what extent English-Dutch and German-Dutch children show evidence of cross-linguistic influence during the real-time processing of long passives and grammatical V2 and ungrammatical V3 structures in Dutch. If a syntactic structure in a bilingual child's one language can activate the same syntactic structure in their other language, as priming studies suggest (e.g., Hsin et al., 2013; Nicoladis, 2012; Serratrice, 2016; Vasilyeva et al., 2012), we expect this co-activation to become visible during sentence processing.

We predict that co-activation is manifested through inhibition effects in our self-paced listening task. Outcomes from online comprehension studies with adult L2 learners suggest that listening to or reading a sentence structure in an L2 which is similar in the L1 can result in less efficient processing (e.g., Foucart & Frenck-Mestre, 2012; Hopp, 2017). We therefore expect English-Dutch children to slow down when listening to V-PP structures in Dutch and German-Dutch children when listening to PP-V and V2 orders. Listening to one of these structures in Dutch will activate the same structure in English/German. We suppose that this will lead to spreading activation in the last two languages. In turn, bilingual children will have to allocate processing resources to inhibit the activation of English/German, which will result in delays during listening.

We further hypothesize that language overlap modulates the relative amount of co-activation of the language not in use and – consequently – the strength of cross-linguistic influence online. To be more precise, we expect online cross-linguistic to be strongest in a situation of *partial overlap*, less strong in a situation of *complete overlap* and absent in a situation of *no overlap*. We explain this in terms of frequency effects, in line with Runnqvist et al. (2013).

In the case of *partial overlap*, structures are more frequent in the language with only one option than in the language with multiple options. Such is the case for the V-PP word order in English and Dutch and the PP-V word order in German and Dutch. As a consequence, we expect these structures to be more strongly associated with and more easily activated in English and German than in Dutch – assuming equal exposure to both languages. If structural co-activation across languages results in delay, *partial overlap* will strengthen this. To be more precise, if the relative frequency of a structure is higher in English or German than in Dutch, processing this structure in Dutch is likely to strongly co-activate the overlapping structure in English/German. In contrast, activation of the structure in Dutch might be relatively weak, as the structure is less frequent. As a consequence, a relatively large amount of processing resources has to be allocated to inhibit co-activation of English/German and to select the Dutch structural representation, resulting in delays online.

In the case of *complete overlap*, in our study the V2 in German and Dutch, frequency distributions are equal across languages. As a consequence, the amount of co-activation German receives while processing the V2 order relative to activation of Dutch will be less compared to a situation of *partial overlap*. Therefore, we predict less strong effects of inhibition and delay for V2 sentences.

In the case of *no overlap*, in our study the PP-V, V2 and V3 word orders in English and Dutch and the V-PP and V3 word orders in German and Dutch, we predict no syntactic co-activation across languages. Because these word orders are absent in Dutch and/or English or German, there is no equivalent word order to prime while processing Dutch sentences. Therefore, we expect no additional co-activation of English or German and, consequently, no inhibition or delay.

Our second research question concerns language dominance and asks to what extent language dominance predicts the occurrence and strength of cross-linguistic influence. We hypothesize that cross-linguistic influence becomes stronger the more dominant children are in English or German relative to Dutch, similar to findings from offline comprehension and production studies with bilingual children (e.g., Bosch & Unsworth, 2020; Foroodi-Nejad & Paradis, 2009) and online studies with adult L2 learners (e.g., Almán Bañón et al., 2018; Hopp, 2017). In line with accounts that consider language co-activation as a possible source of cross-linguistic influence, we expect a greater co-activation of the structural representations of the language not in use when that language is more dominant. As a consequence, slowdown effects should increase the more dominant children are in English or German.

3.3 Method

Participants

Participants were 40 English-Dutch and 42 German-Dutch bilingual children between 5 and 9 years old (English-Dutch: mean age = 7.15, *SD* = 1.37, range = 5.1 – 9.8; German-Dutch: mean age = 7.13, *SD* = 1.48, range = 5.0 – 9.6). They had either acquired both languages from birth (26 English-Dutch bilinguals; 32 German-Dutch bilinguals) or English or German from birth and Dutch before age 4 (14 English-Dutch bilinguals; 10 German-Dutch bilinguals).

Dutch monolingual children (n = 39) served as controls (mean age = 7.26, *SD* = 1.27; range = 5.1 – 9.9). Groups were matched on age ($F(2, 118) = 0.105$; $p > .1$) and socioeconomic status, measured in terms of whether or not parents had finished tertiary education (Fisher's exact test: maternal education level: $p > .05$; paternal education level: $p > .1$). Parents gave written or digital consent.

Information about children's patterns of language history, exposure, and use was collected using a detailed parental questionnaire (Bilingual Language Exposure Calculator; Unsworth, 2013). Children's current relative exposure (*Current input*) as well as their relative exposure over time (*Cumulative input*) to both languages were calculated based on children's language input in different contexts (at home, at school, playing with friends, during holidays, etc.). These two measures served as a proxy of children's language dominance (following e.g., Unsworth, Chondrogianni, & Skarabela, 2018). An overview of background variables is presented in Table 3.1.

Table 3.1. Overview of background variables for bilingual children (means, standard deviations and ranges).

		English-Dutch	German-Dutch	Test statistics[a]
Country of birth	Netherlands (n)	34	41	
Age of first exposure	AoA Dutch > 12 months (n)	4	5	
	Dutch (months)	3.6 (7.60)	4.1 (9.85)	W = 910; p = .422
		0-32	0-36	
	English/German (months)	0	0	
Cumulative input	English/German (%)	48.5 (20.71)	45.4 (19.70)	W = 912; p = .507
		8.1-85.2	15.5-90.9	
Current input	English/German (%)	38.9 (19.18)	35.5 (16.50)	t = .858; p = .394
		9-82	11-72	

[a]t-test when normally distributed, otherwise Wilcoxon-rank sum test.

Materials and design

Self-paced listening task

Children's online sentence comprehension was measured using a self-paced listening task (Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996). In this task, suitable for younger children with little or no reading skills (e.g., Booth, MacWhinney, & Harasaki, 2000; Marinis, 2010), children listen to sentences segment by segment using a button-box.

A total of 30 pairs of test sentences were created: 15 PP-V, 15 V-PP, 15 V2 and 15 V3 structures (Table 3.2, for a complete list of stimuli see Table A3.1 and A3.2 in the appendix). Lexical items used were selected using preschool word lists (Bacchini, Boland, Hulsbeek, Pot, & Smits, 2005; Zink & Lejaegere, 2002). All but one verb in the passive sentences were taken from Armon-Lotem et al. (2016). Sentences were cut constituent-by-constituent rather than word-by-word (see Table 3.2). This way we limited the number of interruptions during listening, keeping sentence processing as natural as possible. The critical region was segments 3 and 4 for the passives and segments 2 and 3 for the V2/V3 sentences, with the spill-over region at segment 5 and segment 4, respectively. All items were recorded by a female native speaker using neutral prosody and intonation and were segmented afterwards. Pictures of animals (without acting out the actions) were shown in a random position on the screen to offer visual support. Comprehension questions were asked after 8 passive and 8 V2/V3 items (equal number of *yes* and *no* responses); these did not query the critical region itself.

Stimuli were distributed over pseudo-randomized lists with each sentence only appearing in one condition. For each child a pair of identical lists was constructed differing only in word order, such that children listened to every sentence in both conditions. The experiment was created in E-Prime, version 2.0 (W. Schneider, Eschmann, & Zuccolotto, 2002).

Before the experimental phase started, children were presented with 10 practice items (5 long passive and 5 V2/V3 clauses), receiving as much feedback as necessary for them to understand the task. In order to move from audio segment to audio segment, they were instructed to press a button on the button box. It was possible for them to move to the next audio fragment before the end of the previous fragment. They could not go back, however. Each experimental list was divided into 5 blocks of 6 items, with breaks throughout when needed.

Table 3.2. Example long passive and V2/V3 sentences in self-paced listening task in PP-V, V-PP, V2 and V3 order.

		Segment						
		1	2	3	4	5	6	7
Long passive	PP-V	de leeuw	wordt	door de beer	gekieteld	en	de muis	likt aan een ijsje
		the lion	*is*	*by the bear*	*tickled*	*and*	*the mouse*	*licks an ice cream*
	V-PP	de leeuw	wordt	gekieteld	door de beer	en	de muis	likt aan een ijsje
		the lion	*is*	*tickeld*	*by the bear*	*and*	*the mouse*	*licks an ice cream*
	Comprehension question: Likt de muis? *licks the mouse* Is the mouse licking?							
Active	V2	op de bank	zingt	de slang	een liedje	van school		
		on the couch	*sings*	*the snake*	*a song*	*from school*		
	V3	*op de bank	de slang	zingt	een liedje	van school		
		on the couch	*the snake*	*sings*	*a song*	*from school*		
	Comprehension question: Is de slang onder de bank? *Is the snake under the couch*							

Sentence repetition tasks
Children's proficiency in Dutch and German or English was measured using sentence repetition tasks (LITMUS-SRep: e.g., Hamann & Abed Ibrahim, 2017; Marinis & Armon-Lotem, 2015). Previous research has shown that, in addition to inevitably involving a memory component, sentence repetition tasks also assess lexical and morphosyntactic language skills (e.g., Polišenská, Chiat, & Roy, 2015). The short version of the Dutch and English tasks consisted of 30 sentences. For German, 30 of the 45 sentences in the original short task were selected to match the Dutch and English sentences in terms of the (difficulty of) structure.

All sentences were recorded by native speakers. For English, a British and an American version were created. Sentences were presented auditorily in a PowerPoint presentation through headphones.

Children received one point for repeating a sentence verbatim, and no points otherwise. In the German task, the chances of making an error were considerably higher compared to the other two languages due to gender and case. Consequently, gender and case errors on German determiners were ignored unless they resulted in a different meaning. Children could receive a maximum of 30 points on each task.

Children's scores on the sentence repetition tasks were used as a third measure of language dominance. More specifically, following Yip and Matthews (2006; see also Unsworth et al., 2018), we calculated relative proficiency scores by subtracting children's score on the English/German task from the Dutch task. A differential score higher than 0 thus meant that children were more proficient in Dutch than English or German, whereas the reverse pattern was reflected in a score lower than 0.

Digit span task
Children's verbal short-term and working memory abilities were assessed using a forward and backward digit span task (Automated Working Memory Assessment (AWMA): Alloway, 2012). The standard scoring procedure of the AWMA was used (forward: max. 48; backward: max. 36).

Wechsler non-verbal intelligence scale
To ensure comparability across groups, nonverbal intelligence score was measured using the Wechsler Nonverbal-NL (Wechsler & Naglieri, 2008). Due to time constraints only two out of four subtasks were conducted, *Matrix reasoning* for 5 to 7 year olds and *Recognizing* for 8-and-9-year-olds. Norm scores were calculated.

Procedure

Children were tested during two test sessions at home or school, approximately one week apart (minimum two days, maximum 3 weeks). The order of tasks was: self-paced listening, Wechsler non-verbal, digit span and Dutch sentence repetition task in session 1, and self-paced listening and English/German sentence repetition task in session 2, before which children watched a short 3-minute movie in English/German to facilitate the language switch.

Data preparation

To establish whether children were paying attention to the self-paced listening task, A-prime scores were calculated for the comprehension questions, with .5 showing chance performance and 1 indicating perfect performance (e.g., Stanislaw & Todorov, 1999). Children with an A-Prime score at or below chance (< .55) were excluded from further analyses (3 English-Dutch children and 1 Dutch monolingual child).

Segments in the self-paced listening task differed in audio length. This is a common issue in self-paced listening and reading tasks (e.g., Chondrogianni, Marinis, et al., 2015; Ferreira & Clifton, 1986). Following standard procedures (e.g., Marinis, 2010), we therefore calculated residual reaction times (RTs) by subtracting the duration of each audio fragment from participants' raw RTs. Residual RTs above 2500 ms and below 300 ms were removed from the data. Because the distribution of the residual RTs was positively skewed, the data were log transformed.[1] Next, data from children and items deviating more than 2.5 standard deviations from the mean of a segment were removed (one monolingual child in the passive condition). Finally, outlier trials, defined as 2.5 SDs above or below the segment mean by group, word order and child, were removed. In total, less than 3.5% of the RTs in the long passives and in the V2/V3 sentences were removed.

3.4 Results

Background measures

Table 3.3 shows the results for the three groups on the background measures. There were no significant differences between groups.

[1] Residual RTs were sometimes negative due to children having pressed the button before the end of the sound fragment. Therefore a constant was added to the residual RTs before applying the log-transformation.

Table 3.3. Scores on background measures for monolingual and bilingual children (means, standard deviations, and ranges).

		2L1		L1	
		English-Dutch	**German-Dutch**	**Dutch**	**Test statistics[a]**
Sentence repetition	Dutch (max. 30)	19.1 (8.0) 2-29	21.6 (6.9) 4-30	21.0 (6.4) 4-30	$F(2,114) = 1.34$; $p = .266$
	English/ German (max. 30)	16.2 (8.0) 2-29	17.6 (7.8) 0-30		$F(1,77) = 0.62$; $p = .432$
	Difference score	2.9 (8.5) -19-20	4.1 (7.7) -19-22		$F(1,77) = 0.40$; $p = .529$
Digit span	Forward	23.5 (2.8) 18-29	21.7 (4.2) 12-30	22.2 (3.5) 14-31	$F(2,114) = 2.50$; $p = .087$
	Backward	9.6 (3.1) 6-19	9.3 (4.5) 4-24	8.8 (3.0) 1-16	$F(2,114) = 0.48$; $p = .622$
Nonverbal IQ	Standard scores	103.4 (9.5) 84-123	102.1 (12.5) 74-128	105.5 (10.4) 81-126	$F(2,114) = 0.94$; $p = .392$

[a]ANOVAs for comparisons between three groups and t-tests for comparisons between two groups.

Self-paced listening task

All data were analysed using multi-level linear effects analysis in R (version 1.0.153, R Core Team, 2018; package lme4, version 1.1-19, Bates, Maechler, Bolker, & Walker, 2015; and lmerTest, version 3.0-1, Kuznetsova, Brockhoff, & Christensen, 2017). An alpha level of .05 was used for all statistical tests. All continuous variables were centered based on their grand mean. All analyses contained random intercepts by *Item* and *Participant*, random slopes by *Word order* and *Segment*, fixed effects of children's *Age*, *RT on the previous trial*, *Trial number*, *Forward digit span* and *Duration of the audio fragment* and the interaction between *Segment* and *Word order*.[2] When a model did not

[2] Even though we subtracted the duration of the audio fragments from children's total RTs, there was still a clear relationship between audio duration and residual RTs: the longer an audio fragment, the more time a child had to prepare the button press, the

converge, random slopes and intercepts were removed until the model did converge. For all models reported, data point had been removed with absolute standardized residuals exceeding 2.5 (which led to less than 3.5% data removal for each model).

To assess effects of cross-linguistic influence we ran models with bilingual and monolingual children's residual RTs on the pre-critical region (long passives: segment 2; V2/V3: segment 1), the critical region (long passives: segment 3 and 4; V2/V3: segment 2 and 3) and the spill-over region (long passives: segment 5; V2/V3: segment 4) as dependent variable. In a first step, we tested the main effect of *Group* (English-Dutch, German-Dutch and Dutch monolingual). In a second step, we tested the 3-way interaction between *Group*, *Word order* (PP-V/V-PP and V2/V3) and *Segment*. A main effect or interaction was deemed significant when it significantly improved a model without the effect or interaction. Significant effects and interactions were explored using model summaries containing treatment contrasts. Models were relevelled when necessary.

To test for effects of language dominance we ran separate models for the two bilingual groups. We used three dominance measures, namely the percentage *Current input* and *Cumulative input* in English/German and *Relative proficiency* (operationalized using SRT differentials). In these analyses we focused on the critical segments. In a first step, 2-way interactions between the three dominance measures and *Word order* were tested and, in a second step, the 3-way interactions between the dominance measures, *Word order* and *Segment*. Because the three dominance measures strongly correlated (English-Dutch group: *Relative proficiency* and *Cumulative input*: $r = -.69$, $p < .001$; *Relative proficiency* and *Current input*: $r = -.80$, $p < .001$; *Current input* and *Cumulative input*: $r = .90$, $p < .001$; German-Dutch group: *Relative proficiency* and *Cumulative input*: $r = -.70$, $p < .001$; *Relative proficiency* and *Current input*: $r = -.71$, $p < .001$; *Current input* and *Cumulative input*: $r = .88$, $p < .001$), separate analyses were run for each. Similar procedures for significance testing and exploration of effects as discussed in the above were applied.

Long passive

Main analyses. Figure 3.1 shows children's average residual RTs in the PP-V and V-PP condition (for the group's average RTs and standard deviations, see Table A3.3 in the appendix). Overall, the bilingual children had smaller residual

---

smaller their residual RTs. Therefore, we decided to remove this variation from the data by adding audio fragment duration as a predictor in the models.

RTs than the monolingual group. Furthermore, the German-Dutch group was slightly faster than the English-Dutch group.

The main effect of *Group* was significant ($X^2 = 8.0$; $\Delta df = 2$; $p = .018$) as well as the 3-way interaction between *Group*, *Word order* and *Segment* ($X^2 = 20.9$; $\Delta df = 6$; $p = .002$). Simple interactions between *Group* and *Word order* in the model summaries showed that the 3-way interaction was caused by the German-Dutch children behaving different from the other two groups at segments 4 and 5 (see Table 3.4). In other words, the effect of *Word order* in the German-Dutch differed significantly from the effect of *Word order* in the other two groups at these segments, but not at segment 2 and 3. There were no significant differences in the effect of *Word order* between the English-Dutch and the Dutch monolingual children.

To investigate what may have caused the differences between the German-Dutch children and the English-Dutch and monolingual children, we compared their residual RTs in the two conditions separately. Because the German-Dutch children were slightly faster overall than the English-Dutch and monolingual children, a direct comparison of residual RTs at segment 4 and 5 was not possible. Therefore, we explored whether the difference in residual RTs between the German-Dutch and the other two groups at pre-critical segment 2 became significantly larger or smaller at segments 4 and 5.

In the PP-V condition, the difference between the monolingual and German-Dutch children was significantly smaller at segment 4 than segment 2 ($B = 0.0186$; $SE = 0.0046$; $t = -4.051$; $p < .001$; see A3.4 in the appendix for the model summary at segment 4). The difference between the English-Dutch and German-Dutch children was in the same direction but was not significant ($B = 0.0083$; $SE = 0.0046$; $t = 1.803$; $p = .071$). At segment 5, the difference in residual RTs in the PP-V condition between the German-Dutch group and the other two groups did not change significantly as compared to segment 2 (monolinguals: $B = 0.0048$; $SE = 0.0045$; $t = 1.063$; $p = .288$; English-Dutch: $B = -0.0016$; $SE = 0.0045$; $t = -0.347$; $p = .728$; see A3.5 in the appendix for the model summary at segment 5).

In the V-PP condition, the difference in residual RTs between the German-Dutch children and the monolingual or English-Dutch children at segment 2 did not change significantly at segment 4 (monolinguals: $B = 0.0037$; $SE = 0.0046$; $t = 0.801$; $p = 0.423$; English-Dutch: $B = -0.0027$; $SE = 0.0046$; $t = -0.593$; $p = .553$). In contrast, at segment 5 the difference between the German-Dutch group and the other two groups was significantly smaller compared to the difference at segment 2 (monolinguals: $B = 0.0169$; $SE = 0.0045$; $t = 3.740$; $p < .001$; English-Dutch: $B = 0.0089$; $SE = 0.0045$; $t = 1.961$; $p = .050$).

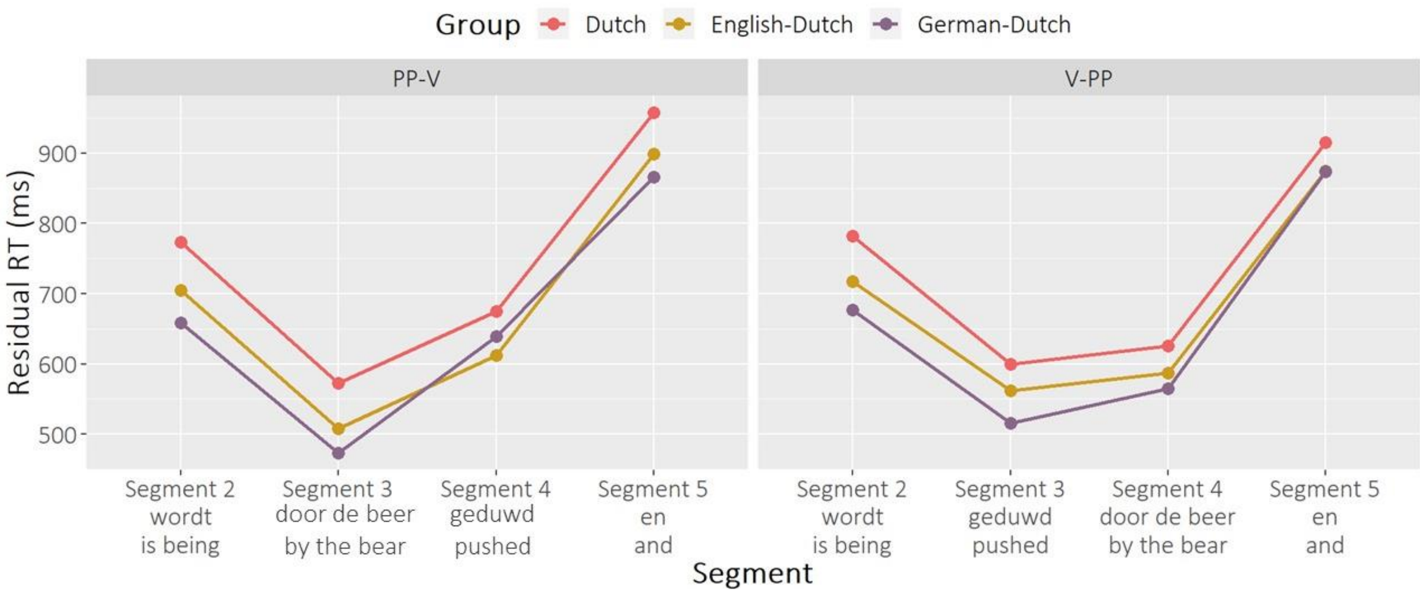


Figure 3.1. Children' s mean residual RTs in the PP-V (left panel) and V-PP (right panel) condition by group on segments 2, 3, 4, and 5.

Table 3.4. Simple interactions between *Group* and *Word order* in the long passive condition at segments 2, 3, 4 and 5. The model was relevelled based on *Group* and *Segment*.

		B	*SE*	*t*	*p*
Segment 2	Dutch monolingual vs. English-Dutch	-0.0030	0.0047	-0.634	.526
wordt *(is being)*	Dutch monolingual vs. German-Dutch	-0.0044	0.0046	-0.947	.344
	English-Dutch vs. German-Dutch	-0.0014	0.0045	-0.306	.760
Segment 3	Dutch monolingual vs. English-Dutch	-0.0030	0.0048	-0.620	.535
door de	Dutch monolingual vs. German-Dutch	0.0005	0.0047	0.101	.920
beer/geduwd *(by the bear/pushed)*	English-Dutch vs. German-Dutch	0.0035	0.0047	0.738	.461
Segment 4	Dutch monolingual vs. English-Dutch	0.0009	0.0049	0.182	.856
geduwd/door de	Dutch monolingual vs. German-Dutch	0.0105	0.0047	2.241	.025
beer *(pushed/by the bear)*	English-Dutch vs. German-Dutch	0.0096	0.0047	2.039	.042
Segment 5	Dutch monolingual vs. English-Dutch	-0.0047	0.0047	-0.997	.319
en *(and)*	Dutch monolingual vs. German-Dutch	-0.0165	0.0046	-3.603	< .001
	English-Dutch vs. German-Dutch	-0.0118	0.0045	2.613	.009

Language dominance. See Table 3.5 for the interactions between the three dominance measures, *Word order* and *Segment* for the English-Dutch and German-Dutch children. In the English-Dutch group, none of the interactions with the dominance measures reached significance. In the German-Dutch group, the 2-way interaction between *Relative proficiency* and *Word order* was significant. None of the other interactions were significant.

Table 3.5. Model improvements after adding interactions between the three dominance measures, *Word order* and *Segment* by group for the long passive sentences.

		English-Dutch			German-Dutch		
		X^2	*Δdf*	*p*	X^2	*Δdf*	*p*
Current input	**Word order*	<0.0	1	.952	0.6	1	.452
	**Word order* Segment*	2.6	2	.272	0.4	2	.826
Cumulative input	**Word order*	0.2	1	.664	3.6	1	.059
	**Word order* Segment*	2.7	2	.257	0.1	2	.935
Relative proficiency	**Word order*	<0.0	1	.975	8.9	1	.003
	**Word order* Segment*	0.9	2	.635	0.1	2	.961

Figure 3.2 shows the interactions of *Relative proficiency* with *Word order* for the German-Dutch group based on the estimated marginal means of the model. German-Dutch children's residual RTs became larger in the PP-V and V-PP conditions the higher their proficiency score was in German relative to Dutch. This slowdown was stronger in the PP-V than in the V-PP condition. This observation was supported by a significant interaction between *Relative proficiency* and *Word order* in the model summaries (B = 0.0009; SE = 0.0003; t = 3.049 p = .004). Children's residual RTs in the PP-V condition were significantly affected by their *Relative proficiency* (B = -0.0018; SE = 0.0008; t = -2.300; p = .027). Simple effects of language dominance were not significant in the V-PP condition (*Relative proficiency*: B = -0.0010; SE = 0.0008; t = -1.236; p = .224).

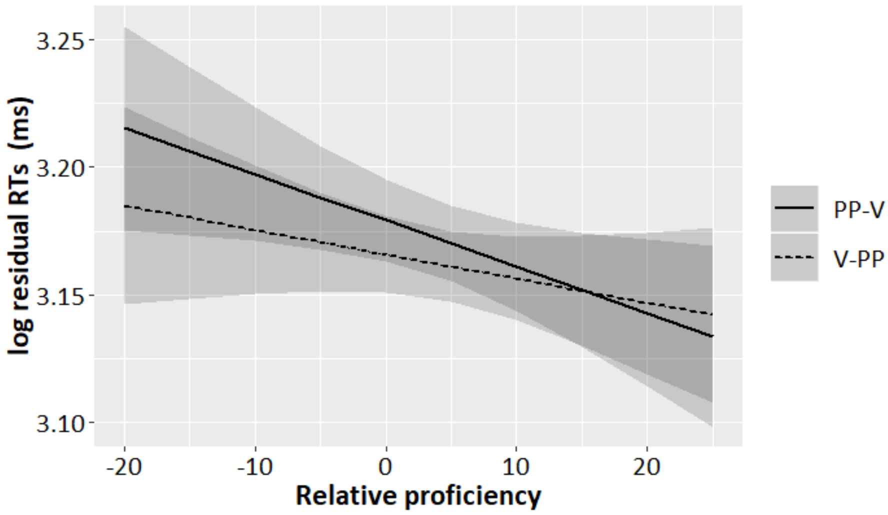


Figure 3.2. Average estimated marginal means of *Relative proficiency* in interaction with the PP-V and V-PP word orders on segments 3, 4 and 5 in the German-Dutch group. A negative relative proficiency score reflects a higher score on the German than on the Dutch sentence repetition task.

In sum, for long passives, German-Dutch bilinguals showed significantly different listening patterns from the English-Dutch and monolingual children at critical segment 4 and spill-over segment 5. This was caused by the German-Dutch group slowing down in the PP-V condition at segment 4 and in the V-PP condition at segment 5 relative to the other two groups. The English-Dutch group showed similar behaviour to the Dutch monolinguals.

German-Dutch children's listening patterns were related to their language dominance profile: the more dominant children were in German – as measured by *Relative proficiency* – the slower they became in the PP-V condition compared to the V-PP condition. No effects of language dominance were found in the English-Dutch group.

V2

Main analyses. Figure 3.3 shows children's average residual RTs in the V2 and V3 word orders (see Table A3.6 in the appendix for children's average RTs and standard deviations on each segment). Again, the bilingual children had smaller residual RTs than the monolingual group, and the German-Dutch group was slightly faster than the English-Dutch group.

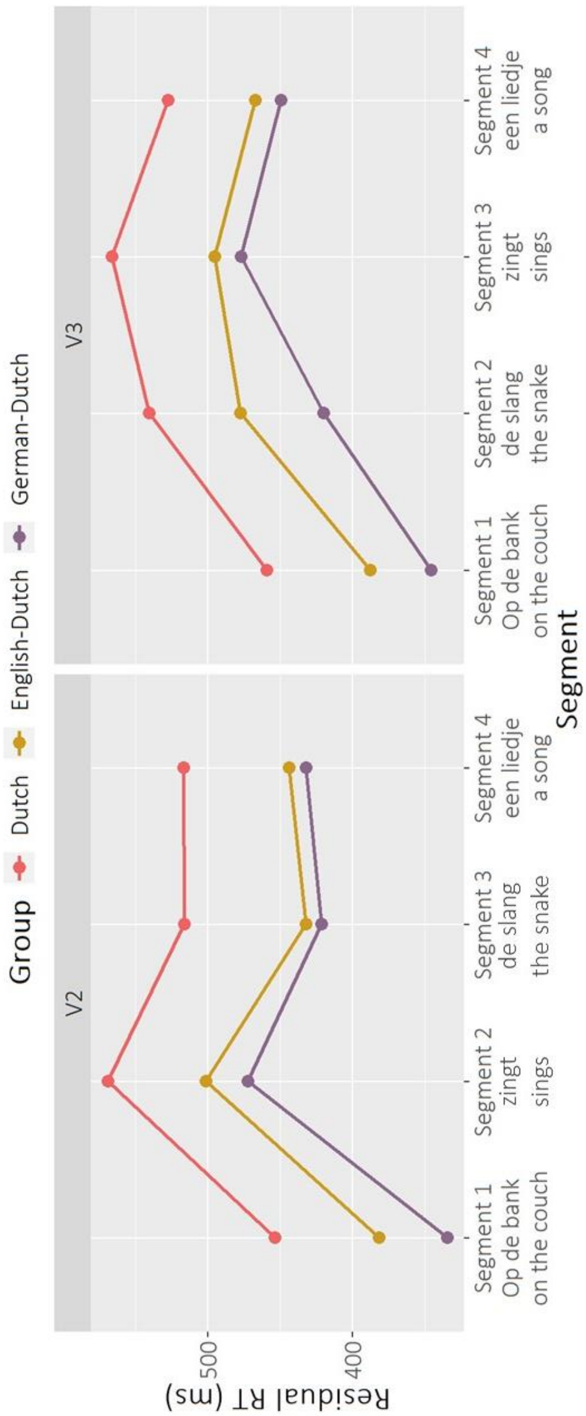


Figure 3.3. Children's mean residual RTs in the V2 (left panel) and V3 (right panel) condition by group on segments 1, 2, 3 and 4.

The main effect of *Group* was significant (X^2 = 12.0; *Δdf* = 2; *p* = .002), whereas the 3-way interaction between *Group*, *Word order* and *Segment* was not (X^2 = 11.7; *Δdf* = 6; *p* = .068). Summary effects showed that the monolingual children had significantly higher residual RTs than the English-Dutch bilinguals (*B* = 0.0247; *SE* = 0.0109; *t* = 2.260; *p* = .026) and the German-Dutch bilinguals (*B* = 0.0359; *SE* = 0.0105; *t* = 3.426; *p* < .001; see A3.7 in the appendix for the summary of the model). The difference in residual RTs between the two bilingual groups was not significant (*B* = 0.0111; *SE* = 0.0108; *t* = 1.034; *p* = .304).

Language dominance. See Table 3.6 for the interactions between the three dominance measures, *Word order* and *Segment* for the English-Dutch and German-Dutch children. In the English-Dutch group none of the 2- or 3-way interactions with our dominance measures were significant. In the German-Dutch group, the 3-way interactions with *Cumulative input* and *Relative proficiency* were significant. The other interactions did not significantly improve the models.

Table 3.6. Model improvements after adding interactions between the three dominance measures, *Word order* and *Segment* by group for the V2/V3 sentences.

		English-Dutch			German-Dutch		
		X^2	*Δdf*	*p*	X^2	*Δdf*	*p*
Current input	*Word order	0.5	1	.469	0.5	1	.461
	Word order Segment	2.7	2	.261	2.8	2	.242
Cumulative input	*Word order	3.8	1	.051	0.4	1	.540
	Word order Segment	2.7	2	.257	8.0	2	.018
Relative proficiency	*Word order	1.0	1	.317	0.1	1	.749
	Word order Segment	0.6	2	.724	9.2	2	.010

Figure 3.4 shows the interaction between *Relative proficiency* and *Word order* by segment (patterns are similar for *Cumulative exposure* and *Word order*).

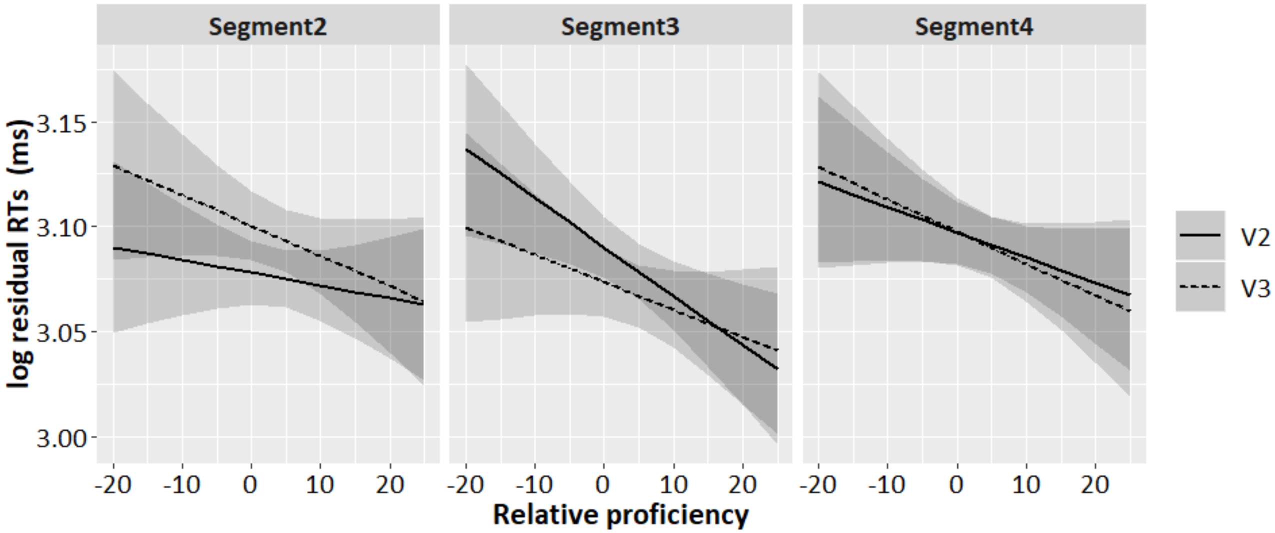


Figure 3.4. Average estimated marginal means of the interaction between Relative proficiency and the V2 and V3 word orders on segments 3, 4 and 5 in the German-Dutch group. A negative relative proficiency score reflects a higher score on the German than on the Dutch sentence repetition task.

At segments 2 and 4, there were no significant differences in the effect of language dominance on children's behaviour in the two word orders (Segment 2: *Cumulative input*: $B = 0.0003$; $SE = 0.0002$; $t = 1.679$; $p = .094$; *Relative proficiency*: $B = -0.0008$; $SE = 0.0005$; $t = -1.821$; $p = .070$; Segment 4: *Cumulative input*: $B = -0.0001$; $SE = 0.0002$; $t = -0.315$; $p = .753$; *Relative proficiency*: $B = -0.0003$; $SE = 0.0005$; $t = -0.706$; $p = .481$).

At segment 3, the greater children's relative exposure to and proficiency in German, the slower they became in the V2 condition compared to the V3 condition, which was significant for both *Cumulative input* ($B = -0.0004$; $SE = 0.0002$; $t = -2.147$; $p = .033$) and *Relative proficiency* ($B = 0.0010$; $SE = 0.0005$; $t = 2.189$; $p = .029$). The simple effect of language dominance was significant only in the V2 condition (*Cumulative input*: $B = 0.0007$; $SE = 0.0003$; $t = 2.093$; $p = .042$; *Relative proficiency*: $B = -0.0023$; $SE = 0.0008$; $t = -2.834$; $p = .007$) and not in the V3 condition (*Cumulative input*: $B = -0.0003$; $SE = 0.0004$; $t = 0.816$; $p = .419$; *Relative proficiency*: $B = -0.0013$; $SE = 0.0009$; $t = -1.431$; $p = .159$).

In summary, at the group level the monolingual and bilingual groups showed similar listening patterns in the V2 and the V3 condition. However, language dominance significantly influenced the listening patterns in the German-Dutch group: at critical segment 3 the more German-dominant children were – as measured by *Cumulative input* and *Relative proficiency* – the more children slowed down in the V2 condition.

3.5 Discussion

In this study, we investigated whether bilingual children show cross-linguistic influence in sentence processing, and whether the presence and strength of any such influence is conditioned by surface overlap and language dominance. A self-paced listening task was conducted with 5- to 9-year-old English-Dutch, German-Dutch, and Dutch monolingual children, while they listened to long passives with a pre-verbal (PP-V) and post-verbal by-phrase (V-PP) and to grammatical verb second (V2) and ungrammatical verb third (V3) word orders in Dutch.

Online cross-linguistic influence and language co-activation

We hypothesized that structures in bilingual children's one language can activate overlapping structures in their other language during sentence processing. We expected that children would have to allocate processing resources to inhibit this co-activation. This, in turn, should be reflected in our self-paced listening task as delays in children's listening times.

Our findings supported this hypothesis. First of all, German-Dutch children slowed down when listening to the overlapping PP-V structure in Dutch. Second, when their dominance profile was taken into account, German-Dutch children also slowed down in the overlapping V2 structure. Our findings are in line with results from online comprehension studies showing that adult L2 learners are less efficient in processing their L2 when structures shared overlap with their L1 (e.g., Foucart & Frenck-Mestre, 2012; Hopp, 2017).

Our results also suggest that cross-linguistic influence attested in previous studies in bilingual children's offline comprehension and production can be explained by syntactic co-activation during sentence processing (e.g., Nicoladis, 2012; Serratrice, 2016). Sentence structures in bilingual children's both languages compete for selection during production and comprehension. When a structure from the language not in use receives sufficient activation, it can be selected over a structure from the language in use when the language not in use is not inhibited sufficiently. These are the situations when syntactic co-activation online should become visible in offline comprehension and production. In contrast, when the language not in use is inhibited sufficiently, it should not affect children's offline choices. This would explain why cross-linguistic influence has not always been attested in offline comprehension and production studies (e.g., Argyri & Sorace, 2007; Mykhaylyk & Ytterstad, 2017).

Our results furthermore suggest that online cross-linguistic influence in bilingual children is modulated not only by syntactic overlap, but by lexical overlap as well. Only the German-Dutch children in our study showed evidence for syntactic co-activation and not the English-Dutch children. We argue that in the first but not in the second group, language co-activation and therefore inhibition was strong enough to become visible in children's listening times. This is because the amount of lexical overlap – as measured by phonological similarity – between Dutch and German is greater than between Dutch and English (Schepens, van der Slik, & van Hout, 2013). This was also the case in the words used in our experiment. First of all, the number of cognates in the self-paced listening task was greater for German than English.[3] Just over two thirds of the words at the critical segments were

[3] Words were considered cognates between languages when there were no more than two sound alternations between words. For the verbs, stems were compared instead of the inflected verbs.

cognates between Dutch and German: 22 of the 30 words (i.e., 15 verbs and 15 nouns) in the long passives and 23 in the V2/V3 sentences. In comparison, only 14 and 16 words were cognates between Dutch and English in the two conditions, respectively. Second, the form of the verbs in our long passive sentences overlap more in Dutch and German than in Dutch and English. To be more precise, the Dutch passive auxiliary *wordt* is similar in form and semantics to the German *wird*, but differs from the English *is being* (e.g., Verhagen, 1992). Furthermore, the morphological construction of past participles is similar in Dutch and German but less similar in Dutch and English (compare *bitten* to Dutch *gebeten* and German *gebissen*). As a consequence, in our experiment German may have been more strongly co-activated at the lexical level than English. If this is correct, the amount of online (lexical) competition from German will have been stronger than that from English. Consequently, German-Dutch children had to spend more processing resources inhibiting German co-activation during sentence processing in Dutch than English-Dutch children had to use to inhibit English co-activation. Additional syntactic co-activation subsequently resulted in visible online cross-linguistic influence in the German-Dutch children. However, in the English-Dutch children, co-activation was not strong enough for visible effects to obtain in listening times.

Our findings are in line with Hopp's (2017) results. He found that in highly proficient L2 speakers, syntactic co-activation of their L1 (German) only became visible in their L2 (English) when the overall co-activation of German was strong enough. This was the case when participants were in a bilingual mode, but not in a monolingual mode. It was only in the bilingual mode that the allocation of processing resources to inhibit structural competition from German slowed down sentence reading in English. Similarly, in our study syntactic co-activation resulted in slowdown effects during listening only when overall co-activation of children's other language was strong enough.

Finally, we obtained one effect in the German-Dutch group that was not predicted based on partial overlap. German-Dutch children slowed down while listening to the V-PP structure at the spill-over segment compared to the other two groups. The V-PP structure does not overlap between Dutch and German, because German does not typically allow material outside an AuxV construction due to its *Satzklammer* structure (e.g., Betz, 2008; Dürscheid, 2012; Haider, 2010). As a consequence, we expected no structural co-activation of German during the processing of Dutch V-PP structures and, in turn, no inhibition effects. It should be noted, however, that AuxVO structures do sometimes appear in spontaneous speech, even though they are not part of the standard grammar (e.g., Betz, 2008; Dürscheid, 2012;

Haider, 2010). This means that (some of) the German-Dutch children in our study may have been exposed to V-PP structures in German and may even have this structure in their German grammar. This is not unlikely, given that most German-speaking parents were L2 speakers of Dutch; thus, they might themselves show cross-linguistic influence from Dutch to German with regard to this structure, which they then also use when addressing their children (also see Paradis & Navarro, 2003 and Sorace et al., 2009, for similar claims about the qualitity of language input bilingual children receive). Consequently, for some of the children at least, the V-PP structure might have been an available structural representation in their German. If this is correct, the slowdown effect found in the German-Dutch children for the V-PP structure could be explained by structural co-activation during sentence processing.

Online cross-linguistic influence and language overlap

With regard to language overlap, we hypothesized that co-activation and consequently cross-linguistic influence would be related to frequency. More specifically, the less frequent a structure is in one of the bilingual children's languages compared to an overlapping structure in the language not in use, the stronger the co-activation of the latter during processing. Such an unequal frequency distribution across languages is present in a situation of *partial overlap*, that is, when the language being processed has more than one option for a certain morphosyntactic property and the language not in use has only one option. Therefore, we predicted stronger effects of online cross-linguistic influence in situations of *partial overlap* than in situations of *complete overlap*. Furthermore, we predicted no online cross-linguistic influence in situations of *no overlap*.

The results from the German-Dutch children corroborated our hypothesis. We found the clearest effect of online cross-linguistic influence in the *partially* overlapping PP-V structure. In particular, German-Dutch children slowed down at the group level when processing long passives. We also observed an effect of online cross-linguistic influence in the *completely* overlapping V2 structure. This effect was however only present when children's dominance profile was taken into account, not at the group level. There was no evidence for cross-linguistic influence in the *non*-overlapping V3 structure.

Our findings are in line with the results of offline comprehension and production studies with bilingual children where cross-linguistic influence has been shown to be likely in situations of *partial overlap* (Foroodi-Nejad & Paradis, 2009; Hulk & Müller, 2000), but not completely ruled out in *no overlap* contexts (Argyri & Sorace, 2007; Nicoladis & Gavrila, 2015). The explanation proposed here for effects of language overlap *online* can also account for differences in findings between overlap situations *offline* and in production. Thus, if syntactic co-activation during sentence processing is largest with *partial overlap*, the chances of cross-linguistic influence becoming visible in offline comprehension and production are high. In *complete* and *no overlap* situations, co-activation can lead to visible effects in children's sentence production or comprehension, but is less likely to be strong enough to become visible. Our findings for language overlap speak against a strong version of Hulk and Müller's (2000; Müller & Hulk, 2001) language overlap hypothesis. Apparently, *partial overlap* is not a necessary condition for cross-linguistic influence to occur online.

The hypotheses tested in this study were in part generated on the basis of sentence processing research in sequential bilingual adults (Foucart & Frenck-Mestre, 2012; Hopp, 2017; Runnqvist et al., 2013). Sentence processing in one language is likely to be less efficient in situations in which bilinguals' languages overlap in word order, but differ in terms of frequency of a certain structure. To our knowledge this study is the first to directly relate the construct of *partial overlap* from the field of child bilingualism to cross-linguistic influence in bilingual sentence processing. In doing so, we were also able to explain why certain sentence structures in adult L2 processing might be more difficult to process than others (e.g., Foucart & Frenck-Mestre). Hence, considering the adult bilingualism literature through the lens of the child bilingualism literature helps to better understand findings for bilingual adults as well.

We argued that the observed slowdown effect for the V-PP structure by the German-Dutch children may result from (some of) the children allowing this order in their German. We based this on the acceptability of this structure in spoken German, and because some of the parents – themselves bilingual in German (L1) and Dutch (L2) – may have shown cross-linguistic influence for this structure. Such an account fits nicely with the findings in Runnqvist et al. (2013). Faster online behaviour of the monolingual children in our study and the preferred production by adult Dutch native speakers of the V-PP structure (e.g., Bernolet, Hartsuiker and Pickering, 2009) suggest that the V-PP structure is more frequent than the PP-V structure in Dutch. Runnqvist and colleagues found clear effects of online cross-linguistic influence from Mandarin only in the more frequent prenominal possessive structure in English. They argued that syntactic co-activation over time in the participants in their study might only have been strong enough to show up in the already preferred prenominal structure and not in the less preferred postnominal structure. This could explain why we found a clear slowdown effect in the V-PP structure in Dutch, even though V-PP is not the canonical word order in German.

Online cross-linguistic influence and language dominance

Our second research question concerned the role of language dominance. We hypothesized that language co-activation in English and German increases the more dominant children are in those languages. We therefore predicted stronger cross-linguistic influence in the form of inhibition the more English- or German-dominant the children were. Language dominance was tested for the two bilingual groups separately and operationalized using three language measures: children's relative current exposure to their languages, their relative cumulative exposure, and their relative proficiency.

The results from the German-Dutch children supported our hypothesis. The more German-dominant bilingual children were, the more they slowed down in the overlapping PP-V and V2 word orders in Dutch. These findings are again in line with a co-activation account of cross-linguistic influence. The more dominant children are in German, the stronger German becomes co-activated when processing sentences with overlapping word order. As a consequence, the more German-dominant children were in our study, the more processing resources they had to allocated to inhibit competition from German when listening to PP-V and V2 structures. This explains why listening times increased in the more German-dominant children while listening to these structures.

Previous studies with bilingual children also obtained stronger effects of cross-linguistic influence in production and offline comprehension the more dominant children were in the language not in use (e.g., Bosch & Unsworth, 2020; Foroodi-Nejad & Paradis, 2009). Similar to language overlap, our explanation of the online role of language dominance helps to account for findings in production and offline comprehension: if syntactic co-activation increases with greater dominance in the language that is not in use, the chances of cross-linguistic influence becoming visible in production and offline comprehension are higher. When the dominant language is the one that is in use, influence from the non-dominant language to the dominant language may still be observed in children's sentence production and interpretations. However, the co-activation of the non-dominant language is less likely to be strong enough to surface than co-activation of the dominant language.

Effects of dominance were most pronounced when dominance was operationalised using children's relative proficiency in German and Dutch. This is in line with studies with adult L2 learners (e.g., Alemán Bañón et al., 2018; Foucart & Frenck-Mestre, 2012; Hopp, 2017), where relative proficiency was also found to relate to patterns of cross-linguistic influence. Similar, although not always significant patterns were found when dominance was operationalised using children's cumulative language exposure. We found no evidence that current input affected the amount of co-activation of German in the German-Dutch children, however. It is unclear why exactly this should be the case. In general, the relationship between language dominance, exposure and proficiency is complex and subject to considerable discussion (e.g., Silva-Corvalán & Treffers-Daller, 2015; Unsworth et al., 2018). Further research is needed to explore these relations in more detail.

3.6 Conclusions

Our study is among the first to investigate cross-linguistic influence in bilingual children's listening times during real-time sentence processing and the first to use the self-paced listening paradigm for this goal. Online cross-linguistic influence was manifested as inhibition during listening when structures were shared between languages. Crucially, cross-linguistic influence in our study was modulated by lexical overlap, structural overlap, and language dominance. Cross-linguistic influence was only attested from German to Dutch, two highly related languages in terms of lexical overlap, and not from English to Dutch. It was stronger in a situation of *partial overlap* rather than in a situation of *complete overlap*, and it became stronger the more dominant children were in German. We argued that these three factors affected the amount of language co-activation during sentence processing and concomitantly the level of inhibition needed to process a unilingual sentence. It was only when sufficient co-activation was present that inhibition became visible in children's listening times.

In conclusion, the use of an online research technique, self-paced listening, allowed us to reveal subtle effects of cross-linguistic influence on listening times during real-time sentence processing in bilingual children. We believe that online studies like this one are crucial if we are to develop a more comprehensive account of cross-linguistic influence in bilingual language development and a better understanding of the processing mechanisms that underpin it.

CHAPTER 4

Cross-linguistic influence in simultaneous bilingual adults' online processing of long passives and Verb Second

A self-paced listening study

Abstract
A self-paced listening study was conducted with simultaneous bilingual adults and adolescent to directly compare cross-linguistic influence during their sentence processing with that in simultaneous bilingual children (van Dijk, Dijkstra & Unsworth, Chapter 3). As far as we know, this is the first study to make such comparison in this way. We collected listening times from English-Dutch and German-Dutch simultaneous bilinguals, and Dutch-monolingually-raised adults and adolescents. Studied structures were Dutch long passives, Verb Second and ungrammatical Verb Third word orders in Dutch. Online cross-linguistic influence manifested itself as a slowdown effect moderated by lexical overlap, surface overlap, language dominance and language mode. In particular, cross-linguistic influence appeared only in the German-Dutch participants, that is in a situation of *partial overlap*, in relation to language dominance and in a bilingual language mode. Cross-linguistic influence in adults and adolescents was qualitatively similar to that in children, but less pronounced. We explain these similarities and differences in terms of syntactic co-activation and inhibition. Crucially, our findings suggest that the same mechanisms underlie cross-linguistic influence during sentence processing in child and adult populations.

Based on: van Dijk, C.N., Dijkstra, T., & Unsworth, S. (unpublished manuscript). Cross-linguistic influence in bilingual adults: a self-paced listening experiment.

4.1 Introduction

How the use of morphosyntax in one language is affected by knowledge of another language has been a prominent topic of research in both bilingual adults and children. For instance, bilingual adults and children have been found to overaccept ungrammatical bare noun phrases in Italian in contexts in which their other language allows such structures (see (1) taken from Kupisch, 2012, p. 746; e.g., Kupisch, 2012; Serratrice, Sorace, Filiaci, & Baldo, 2009).

(1) *Davvero non lo sapevi? Patate crescono sotto terra.
really not it knew.you potatoes grow under earth
"Really, you didn't know that? Potatoes grow under the ground."

Findings in both populations suggest that such *cross-linguistic influence* occurs under specific circumstances (e.g., Hopp, 2010; Runnqvist, Gollan, Costa, & Ferreira, 2013; Serratrice, 2013; van Dijk et al., Chapter 3).

Recent theories argue that similar mechanisms underlie cross-linguistic influence in both adults and children (e.g., Hopp, 2017; Kupisch, 2014; Nicoladis, 2006, 2012; Runnqvist et al., 2013; Serratrice, 2016). One proposed mechanism is *language co-activation*: the processing of morphosyntactic structures in one language might activate similar structures in bilinguals' other language competing for selection (e.g., Hopp, 2017; Kupisch, 2014; Nicoladis, 2006, 2012). This, in turn, necessitates the allocation of processing resources to inhibit the competition from the non-target language. The stronger the co-activation is, the greater the chances are of observing cross-linguistic influence (e.g., Hopp, 2017). Language proficiency, a proxy for language dominance, and language mode have been argued to influence the strength of such co-activation (e.g., Hopp, 2017; Kootstra & Doedens, 2016; van Dijk et al., Chapter 3). Another proposed mechanism is *syntactic priming*. The use of a morphosyntactic structure in one language activates the same structure in the other language (e.g., Hartsuiker, Pickering, & Veltkamp, 2004; Loebell & Bock, 2003). This might then make the structure more readily available for subsequent use in both languages (e.g., Hartsuiker & Bernolet, 2017; Runnqvist et al., 2013; Serratrice, 2016).

However, a comparison of the evidence for adult and child populations is difficult, because their performance has been investigated by groups of researchers working relatively independently. Furthermore, any such comparison is complicated by the different research focus in the two groups. Research with adults has focussed on studying adult second language (L2) learners, i.e., language users who have acquired one language later in life and typically use this language less frequently than their first language (e.g., Clahsen & Felser, 2006; Kroll & Dussias, 2013). In contrast, research with children often examines the behaviour of bilingual children who are acquiring two languages simultaneously (see Serratrice, 2013 for an overview). Yet another difference between the two fields of research lies in whether the studies involved usually employ online or offline techniques. Research with bilingual adults has used both online and offline research techniques, while online research with bilingual children is scarce.

Due to these differences in research traditions, an important issue remains understudied: how does the bilingual processing system in children develop into that of the adolescents and adults they grow up to be later in life? The few studies that have directly compared language processing in children and adults have typically focussed on child and adult second language (L2) learners and employed offline techniques (e.g., Chondrogianni, 2008; Marinis & Chondrogianni, 2011; Unsworth, 2005). Furthermore, cross-linguistic influence was typically not the focus of such studies (but cf. Unsworth, 2005). In the present study, we set out to explore cross-linguistic influence *during* sentence processing in simultaneous language acquisition by comparing the performance of simultaneous bilingual adults, adolescents and children with comparable background profiles on the same online task and materials, and applying the same set of predictor variables.

In an earlier paper (van Dijk et al., Chapter 3), we examined the performance of bilingual children in a self-paced listening task on two syntactic structures in Dutch (discussed in more detail below): long passives and Verb Second (V2). Participants were English-Dutch and German-Dutch bilingual children who started to acquire both of their languages before the age of four and were between five and nine years old at the time of testing. The present study targets participants that can be considered as belonging to the same populations but are older, namely adults and adolescents who have continuously been exposed to both of their languages (English and Dutch or German and Dutch) since the onset of acquisition and who were still using both frequently at the time of testing. We asked these bilingual adults and adolescents to perform the same task as the children in our earlier study. We considered the role of various predictor variables of cross-linguistic influence previously identified in research on bilingual children. Several studies have independently suggested the presence of cross-linguistic influence in both adults and children, but in the present study we were able to compare the effects for the two groups more directly.

To set the stage for our study, we will review the available research on cross-linguistic influence in simultaneous bilingual adults and compare these results to findings from studies with simultaneous bilingual children. Because very little is known about cross-linguistic influence during sentence processing in either of these groups, we first discuss findings from offline comprehension and (elicited) production studies and then turn to online studies. Relatively few studies have investigated cross-linguistic influence in adults that can be considered simultaneous bilinguals (i.e., with the age of onset of acquisition (AoA) before the age of 4;0, e.g., Genesee, Paradis, & Crago, 2004; McLaughlin, 1978; Unsworth, 2013). Furthermore, not every study on this topic reported their participants' exact AoA. We therefore included studies in our literature review involving participants better characterized as early sequential bilinguals (e.g., Unsworth, 2005). We will refer to participants in previous studies as early bilingual adults, as opposed to simultaneous bilinguals.

Cross-linguistic influence in early bilingual adults

Cross-linguistic influence has been observed in early bilingual adults for various language combinations and for different morphosyntactic properties, such as adjective-noun orders (e.g., Kupisch, 2014), specific and generic sentences (e.g., Kupisch, 2012; Kupisch & Barton, 2013; Montrul & Ionin, 2010), *that*-traces (e.g., Martohardjono, Phillips, Madsen II, & Schwartz, 2017), clitics (e.g., Montrul, 2010), and possessive structures (e.g., Anderssen, Lundquist, & Westergaard, 2018; Runnqvist et al., 2013). There is evidence that cross-linguistic influence can manifest itself in two opposing ways. The first manifestation is cross-linguistic influence in the 'classic' sense, that is, in a few studies cross-linguistic influence has been found to result in converging behaviour in bilingual's languages (e.g., Kupisch, 2012; Montrul & Ionin, 2010). For example, Kupisch (2012) found that German-Italian speakers overaccepted ungrammatical bare noun structures in Italian, because such structures are grammatical in German (see (1) above). In other words, bilingual adults' judgements of Italian clauses converged towards their German language system. We refer to such outcomes of cross-linguistic influence as cross-linguistic convergence.

The second manifestation of cross-linguistic influence is more indirect. Adults may try to make their languages as distinct as possible by overproducing or overaccepting morphosyntactic properties in one of their languages that are different from their other language (e.g., Anderssen et al., 2018; Kupisch, 2014; Kupisch & Barton, 2013). For instance, Kupisch (2014) investigated German-Italian bilinguals' judgements and corrections of adjective-noun orders in Italian. In Italian, attributive adjectives most frequently appear in postnominal order (e.g., *una macchina bella*/a car nice, "a nice car", p. 223). However, adjectives can also appear in prenominal order (e.g., *una bella macchina*/a nice car, p. 223). In German, attributive adjectives are prenominal only (e.g., *ein nettes Auto*/a nice car). Kupisch found that German-dominant bilinguals over-accepted incorrect postnominal adjectives and over-corrected correct prenominal adjectives in Italian. In other words, the adults preferred the more frequent postnominal adjective position in Italian as opposed to the canonical German prenominal adjective position. We refer to this (indirect) type of cross-linguistic influence as overcorrection (e.g., Kupisch, 2014).

Surface overlap

Whilst cross-linguistic influence has been attested for a number of morphosyntactic properties, not every morphosyntactic property or every bilingual adult seems equally vulnerable to cross-linguistic influence (e.g., Kupisch et al., 2014; Rinke & Flores, 2014). For example, Rinke and Flores (2014) found no clear evidence for influence of German, a language without clitics, on German-Portuguese speakers' abilities to judge grammatical and ungrammatical Portuguese sentences concerning clitic use and placement. The likelihood of cross-linguistic influence may depend on the type of surface overlap between a bilingual's languages. It has been argued that (partially) overlapping structures are especially sensitive to cross-linguistic influence (e.g., Hulk & Müller, 2000; Müller & Hulk, 2001). In particular, a language with only one option for a morphosyntactic property might reinforce the same option in a language with multiple options for the same morphosyntactic property.

There is indeed some evidence for cross-linguistic influence in situations of *partial overlap* in bilingual adults (e.g., Anderssen et al., 2018; Kupisch, 2014). However, cross-linguistic influence has also been found in situations without language overlap (*no overlap*, e.g., Kupisch & Barton, 2013). Furthermore, considerable variation in the effect of cross-linguistic influence can be observed between participants for the same morphosyntactic property within the same study, even with *partial overlap* (e.g., Andersson et al., 2018; Kupisch, 2012). This implies that surface overlap between languages cannot be the only explanation of cross-linguistic influence.

Language dominance

In addition to surface overlap, a second predictor is language dominance. To date, language dominance in studies on cross-linguistic influence has typically been operationalized in a categorical way by dividing participants into groups based on whether the language tested is spoken in the society they lived and/or grew up in (the majority language) or at home while growing up (the minority language). In direct comparisons, studies found cross-linguistic influence in minority-language speakers but not in majority-language speakers (e.g., Kupisch, 2012; 2014; Kupisch & Barton, 2013). This suggests that cross-linguistic influence is asymmetric and is especially strong in the direction from the bilingual adults' dominant language to their non-dominant language.

Cross-linguistic influence in bilingual children

Similar to studies with early bilingual adults, cross-linguistic influence has been attested in bilingual children for a wide range of language combinations and morphosyntactic properties (see van Dijk, van Wonderen, Koutamanis, Kootstra, Dijkstra & Unsworth, Chapter 2 of this thesis, for an overview). There is ample evidence for cross-linguistic convergence in bilingual children's languages for a large number of morpho-syntactic properties (e.g., van Dijk et al., Chapter 2). As for bilingual adults, cross-linguistic influence has been observed in bilingual children for possessive structures (e.g., Nicoladis, 2002), adjective-noun orders (e.g., Cuza & Pérez-Tattam, 2016; Nicoladis, 2006) and with generic and specific sentences (e.g., Serratrice et al., 2009) as well. Effects in bilingual children's clitic placement are mixed, however, with some studies finding significant effects (e.g., Pérez-Leroux, Cuza, & Thomas, 2011), whereas others did not (e.g., Argyri & Sorace, 2007). Furthermore, to our knowledge there is no evidence for cross-linguistic influence in children's judgments of *that*-traces (e.g., Gathercole, Laporte, & Thomas, 2005).

There is little evidence for overcorrection in bilingual children. On the contrary, with respect to some of the overcorrected linguistic properties in adult bilinguals, cross-linguistic convergence was found in bilingual children, such as adjective-noun orders (e.g., Cuza & Pérez-Tattam, 2016; Nicoladis, 2006) and possessives (Nicoladis, 2002). For instance, whilst English-dominant bilingual adults overcorrected grammatical prenominal Adj-N orders in French compared to French-dominant bilinguals, English-French bilingual children overproduced such prenominal orders compared to French monolingual peers. At the same time, Nicoladis' (2006) study offers some evidence for overcorrection in bilingual children: French-English children also overproduced ungrammatical postnominal N-Adj orders in French.

One of the very few studies providing evidence of overproduction in bilingual children is by Döpke (1998). She examined English-German bilingual's acquisition of verb placement and inflections in German and English. In German, the verb should be placed in second position in main clauses and in final position in subordinate clauses. In contrast, English has a rather rigid SVX order that is maintained regardless of whether a constituent proceeds the subject (i.e., XSVX). Döpke found that bilingual children sometimes spontaneously produced utterances in German with the verb in third position (see the example in 2), something that is unattested in the speech of monolingual peers. This was taken as evidence for cross-linguistic influence from English to German. Döpke also found that – as illustrated in (2) – bilingual children tended to incorrectly mark the main verb in complex verb utterances (i.e., auxXV(X)). This behaviour was inconsistent with influence from English. Rather, Döpke argued that when word order between German and English overlapped, bilingual children were trying to differentiate between their languages by using German verb morphology, even though the verb morphology was incorrect. Hence, children were trying to overcorrect their use of verb inflection in German in order to make their languages contrast, in line with what has since been found for adult bilinguals for various word order patterns (e.g., Kupisch, 2014; Kupisch & Barton, 2013).

(2) Er kann nicht kommt rein.
he can not comes in
"He cannot come in."

Surface overlap

Effects of surface overlap have been extensively studied in bilingual children. Similar to studies with early bilingual adults, cross-linguistic influence has often been observed in children in situations of *partial overlap* (e.g., Hulk & Müller, 2000; Meroni, Smeets, & Unsworth, 2017; Nicoladis, 2006). However, again in line with the adult literature, cross-linguistic influence is not limited to partial overlap situations, but occurs with *no overlap* as well (e.g., Argyri & Sorace, 2007; Foroodi-Nejad & Paradis, 2009; Nicoladis, 2003; van Dijk et al., Chapter 2).

Language dominance

Some studies have found that cross-linguistic influence is seen only in bilingual children's non-dominant language and not in their dominant language (e.g., Argyri & Sorace, 2007; Sorace, Serratrice, Filiaci, & Baldo, 2009), in line with findings from offline studies with bilingual adults (e.g., Kupisch, 2012; 2014; Kupisch & Barton, 2013). However, in contrast to most adult literature, there is ample evidence of cross-linguistic influence in children's dominant language as well (e.g., Foroodi-Nejad & Paradis, 2009; van Dijk et al., Chapter 2). Continuous measures of language dominance can account for individual differences between children. In particular, in some studies the more dominant children were in the language being tested, the smaller effects of cross-linguistic influence became (e.g., Foroodi-Nejad & Paradis, 2009; Kidd, Chan, & Chiu, 2015; Nicoladis, 2006). As a consequence, language dominance can perhaps be better characterized as a factor that predicts the strength rather than the direction of cross-linguistic influence in bilingual children (also see van Dijk et al., Chapter 2). Because language dominance in early bilingual adults is typically operationalized as a categorical predictor, it is unclear whether the same holds for this population (but see Martohardjono et al., 2017, discussed below).

General bilingualism effect

An alternative explanation for the linguistic performance in bilingual adults and children attributed to cross-linguistic influence is a general bilingualism effect (e.g., Sorace, 2011; Sorace & Serratrice, 2009). This term refers to the frequent observation that adult second language learners process their second language less efficiently than native speakers, irrespective of the properties of their first language (e.g., Clahsen & Felser, 2006; Hopp, 2010; Sorace, 2011). Sorace and Serratrice (2009), amongst others (e.g., Sorace et al., 2009), have suggested that a similar effect might be at play in bilingual acquisition. One possible reason is that bilinguals have to suppress the activation of the non-target language during language processing. As a consequence, fewer processing resources might be left for efficient sentence processing. However, as yet, the evidence for a general bilingualism effect in simultaneous bilingual adults or children remains limited (but cf. Sorace et al., 2009).

To sum up, there are some common findings in adult and child studies when it comes to cross-linguistic influence. First, there is evidence that cross-linguistic influence occurs for the same morphosyntactic properties in both bilingual children and early bilingual adults. Second, surface overlap and language dominance have been identified as its predictors in both bilingual populations. At the same time, however, there are also a number of findings which differ across the two populations. First, whilst cross-linguistic influence often shows itself as overcorrection in early bilingual adults, it usually manifests itself as convergence in bilingual children. Second, effects of cross-linguistic influence in early bilingual adults seem limited to their non-dominant language, whereas in bilingual children, cross-linguistic influence has been frequently observed in their dominant language as well. Finally, it is unclear whether (some) observed effects of cross-linguistic influence in adults and children can be explained by a general bilingualism effect instead.

Disadvantages of offline tasks

It is unclear what accounts for the observed differences between adult and child studies. One possible explanation is that cross-linguistic influence is qualitatively different after language acquisition has taken place (i.e., overcorrection in adults and convergence in children). However, recent theories suggest that the same mechanism underlies cross-linguistic influence in bilingual adults and children (e.g., Kupisch, 2014; Nicoladis, 2012; Runnqvist et al., 2013; Serratrice, 2016). As a more likely explanation, we therefore propose that observed differences between populations might be ascribed to methodological differences: the use of offline rather than online tasks.

Specifically, offline tasks might tap into different types of knowledge in adults and children and might give adults an advantage. There are several reasons to believe that this is the case. First, offline tasks allow for the involvement of explicit knowledge (e.g., Marinis, 2010). Given that adults have better developed meta-linguistic skills than children (e.g., Edwards & Kirkpatrick, 1999), it is likely that explicit knowledge plays a larger role in their offline responses than in children.

Second, adult participants were probably well aware that they were tested because of the language combination they grew up with, and they may also well be conscious that the structures tested were to some extent different or similar between their languages. This awareness, in combination with adults' well-developed meta-linguistic skills, might explain their tendency to overcorrect morphosyntactic structures in their languages. Although we do not rule out that children may make use of explicit knowledge in offline tasks as well, we believe that the explicit component in their responses is likely less strong than in adults.

Third, children have less working memory resources available than adults (see Schneider, 2015 for an overview). Because offline tasks place a burden on participants' working memory (e.g., Marinis, 2010), less capacity might remain for engaging meta-linguistic skills during an offline task in bilingual children compared to adults. Hence, findings of cross-linguistic influence might reflect more implicit processes in children and more explicit processes in adults. Furthermore, children might have experienced more difficulties in general performing experiments on cross-linguistic influence than adults.

Fourth, adult bilinguals might be more capable of suppressing effects of co-activation and priming during sentence processing in their offline responses than bilingual children. As discussed above, sentence processing in one language involves the priming and co-activation of similar sentence structures in the language not in use (e.g., Hartsuiker & Bernolet, 2017; Nicoladis, 2006, 2012). This, in turn, might require the allocation of processing skills to inhibit cross-language competition (e.g., Hopp, 2017). When such inhibition fails, cross-language competition can become visible in bilinguals' production and offline judgements and interpretations (e.g., Nicoladis, 2006; 2012; van Dijk et al., Chapter 3). Given that children possess less processing capacity than adults, such offline cross-linguistic influence would be more likely to surface in children than adults. Furthermore, this asymmetry between groups might be increased by the strong involvement of working memory in offline tasks (e.g., Marinis, 2010). Hence, the use of offline tasks makes it difficult to directly compare effects of cross-linguistic influence in early bilingual adults and bilingual children.

Online cross-linguistic influence in early bilingual adults and children

Given problems with offline measures, online measures might be better suited to compare cross-linguistic influence in early bilingual adults and bilingual children. First of all, they tap more into implicit than explicit knowledge (e.g., Marinis, 2010). Second, they are a more direct measure of language co-activation during sentence processing (e.g., van Dijk et al., Chapter 3). And third, they require less involvement of working memory (e.g., Marinis, 2010). Unfortunately, however, there are only a few online studies that have investigated cross-linguistic influence in early bilinguals (e.g., Kupisch, 2012; Lemmerth & Hopp, 2019; Martohardjono et al., 2017; Runnqvist et al., 2013; van Dijk et al., Chapter 3).

The tasks used in online studies with early bilingual adults include timed acceptability/grammaticality judgements (e.g., Kupisch, 2012; 2014), onset of speech production (e.g., Runnqvist et al., 2013), and event-related potentials (ERPs; Martohardjono et al., 2017). Results suggest that cross-linguistic influence can occur during real-time sentence processing. Moreover, Martohardjono and colleagues (2017) showed that online effects may not necessarily emerge offline (i.e., grammaticality judgements), in line with our suggestion that online tasks might be a more direct measure of cross-linguistic influence. Online cross-linguistic influence has been attested in situations of *partial* (e.g., Runnqvist et al., 2013) and *no overlap* (e.g., Martohardjono et al., 2017). Furthermore, it has also been found in participants' dominant language (e.g., Runnqvist et al., 2013). Finally, language dominance – as measured by the amount of exposure to the language not in use – has been observed to predict the strength of cross-linguistic influence online (e.g., Martohardjono et al., 2017). In particular, the less dominant participants were in the language tested, the more cross-linguistic influence they showed. In short, results from online studies with early adult bilinguals show that online measures can reveal subtle effects of cross-linguistic influence that are not necessarily visible in offline comprehension and production.

To the best of our knowledge, only two studies have investigated cross-linguistic influence during sentence processing in bilingual children (i.e., Lemmerth & Hopp, 2019; van Dijk et al., Chapter 3). Lemmerth and Hopp (2019) examined Russian-German children's processing of gender cues using an eye-tracking task in the visual world paradigm and found no evidence of cross-linguistic influence. However, they only investigated situations of *complete* and *no overlap*. Furthermore, they did not take into account children's dominance profiles. In our self-paced listening study (van Dijk et al., Chapter 3), we did find evidence for cross-linguistic influence in German-Dutch children in situations where structures in Dutch partially or completely overlapped with structures in German (discussed in more detail below). The strength of cross-linguistic influence was modified by language dominance. However, we found no evidence of online cross-linguistic influence in a group of English-Dutch children. We explained this in terms of lexical overlap: the greater degree of lexical overlap between German and Dutch as compared with English and Dutch (e.g., Schepens et al., 2013) increased the chances of cross-linguistic influence occurring in German-Dutch children as compared with in English-Dutch children.

Together, these findings indicate that online cross-linguistic influence is present in early bilingual adults and bilingual children. In bilingual children, surface overlap, lexical overlap and language dominance all seem to affect the strength of cross-linguistic influence during sentence processing. Similar results have been obtained for language dominance in adults. It is unclear, however, what the exact role is of surface and lexical overlap during sentence processing in adults, because these overlap types have not been systematically investigated within the same study. Furthermore, it is difficult to directly compare results from online adult and child studies, because studies differ in the methods employed and morphosyntactic properties investigated (e.g., Lemmerth & Hopp, 2019; Martohardjono et al., 2017; Runnqvist et al., 2013; van Dijk et al., Chapter 3). Consequently, it is as yet unknown to what extent effects of cross-linguistic influence are comparable across the two populations.

4.2 The present study

To allow a direct comparison of effects of cross-linguistic influence in bilingual adults and children, we replicated the Dutch self-paced listening experiment we conducted with English-Dutch and German-Dutch bilingual children and Dutch monolingual children (e.g., van Dijk et al., Chapter 3) with English-Dutch and German-Dutch bilingual and monolingually raised Dutch-speaking adults and adolescents. Furthermore, in order to increase the co-activation of English and German, respectively, we replicated the experiment in a monolingual and a bilingual mode. In the monolingual mode, we added Dutch fillers to the Dutch experiment. In the bilingual mode, we added either English (for the English-Dutch group) or German fillers (for the German-Dutch group) to the Dutch experiment. Our aims were to investigate (i) whether effects of cross-linguistic influence during sentence processing are similar in bilingual children and adults; (ii) whether the same predictors found to modulate cross-linguistic influence in our study with bilingual children – i.e., lexical overlap, surface overlap, and language dominance – also do so in adult bilinguals; and (iii) whether bilingual mode modulates cross-linguistic influence.

In the remainder of this section, we discuss the structures of interest – i.e., long passives and V2 sentences – as well as the results from our child study in more detail. Subsequently, we formulate our hypotheses and predictions for the adult situation.

Long passives

Dutch long passives can have a preverbal (PP-V) or postverbal (V-PP) by-phrase (see Table 4.1; Koster, 1974). Studies with adult native speakers of Dutch and Dutch monolingual children suggest that the V-PP order is the preferred one (e.g., Bernolet, Hartsuiker, & Pickering, 2009; van Dijk et al., Chapter 3). Due to its rather rigid SVO order, English only allows the V-PP order (see Table 4.1; e.g., Lehmann, 1978). In contrast, German syntax requires the by-phrase to precede the main verb due to its AuxXV order (PP-V; see Table 4.1; e.g., Dürscheid, 2012). However, the V-PP word order is not entirely ruled out in German as movement to the right side of the verb is sometimes observed (e.g., Betz, 2008; Dürscheid, 2012; Haider, 2010). Hence, there is partial overlap between the V-PP structures in Dutch and English, with the V-PP structure being the only and therefore the more frequent option in English, and between the PP-V structure in Dutch and German, with the PP-V structure in German being – in theory – the only and therefore more frequent option. However, if the V-PP structure sometimes occurs in German as well, there might also be partial overlap between the Dutch and German V-PP structure. Crucially, we expect preferences in Dutch and German to be reversed compared to the PP-V structure. To be more precise, the V-PP structure should be more preferred in Dutch than in German and the PP-V structure should be more preferred in German than in Dutch.

Table 4.1. Placement of the by-phrase in long passives in Dutch, English and German.

	PP-V
Dutch	De beer wordt door de leeuw geduwd. the bear is being by the lion pushed
English	-
German	Der Bär wird vom Löwen geschoben. the bear is being by the lion pushed
	V-PP
Dutch	De beer wordt geduwd door de leeuw. the bear is being pushed by the lion
English	The bear is pushed by the lion.
German	?Der Bär wird geschoben vom Löwen. the bear is being pushed by the lion

V2

With regard to V2 sentences, Dutch and German always require the verb to be in second position in main clauses (e.g., Haider, 2010; Koster, 1975; Zwart, 2011). Therefore, subject-verb inversion takes place in sentences initiated by an adverb (see Table 4.2). In English, however, the verb should follow the subject in main clauses regardless of sentence initial material (see Table 4.2; e.g., Lehmann, 1978), with the exception of a few structures (e.g., *When did she eat an apple?*; Radford, 2004). Hence, V2 constitutes a situation of *complete overlap* between Dutch and German and a situation of *no overlap* between Dutch and English. Furthermore, Verb Third (V3) orders constitute a situation of *no overlap* between Dutch and English and Dutch and German.

Our results with bilingual children showed that German-Dutch children slowed down when they were listening to sentences that overlapped in word order between German and Dutch. These effects of online cross-linguistic influence were modulated by surface overlap and language dominance. Thus, slowdown effects were most pronounced in partially overlapping sentences (i.e., PP-V and V-PP structures) and less so in completely overlapping sentences (i.e., V2 structures). Language dominance further affected the strength of cross-linguistic influence: the more German-dominant children were, the more they slowed down in the PP-V and V2 word orders.

Table 4.2. Verb placement in adverb-initial clauses in Dutch, English and German.

	V2
Dutch	Op de bank zingt de slang een lied. on the couch sings the snake a song
English	-
German	Auf der Sofa singt die Schlange ein Lied. on the couch sings the snake a song
	V3
Dutch	*Op de bank de slang zingt een lied. on the couch the snake sings a song
English	On the couch the snake sings a song.
German	-

Hypotheses

Cross-linguistic influence and language co-activation
Our first hypothesis was that if language co-activation and priming underlie online cross-linguistic influence in both adults and children (e.g., Nicoladis, 2006; 2012; Serratrice, 2016), similar effects should be observed in both populations. We therefore predicted that the simultaneous bilingual adults and adolescents in our study would also slow down when listening to sentences with a word order overlapping between their languages.

The role of lexical overlap, surface overlap, language dominance and mode
Our second hypothesis was that lexical overlap, surface overlap, language dominance and language mode would modulate the strength of cross-linguistic influence in simultaneous bilingual adults (e.g., Argyri & Sorace, 2007; Hulk & Müller, 2000; van Dijk et al., Chapter 3).

With regard to lexical overlap, and in line with our study with bilingual children, we hypothesized that cross-linguistic influence will become stronger with increasing lexical overlap between the language not in use and the language being processed. In particular, German shares more lexical overlap with Dutch than English (e.g., Schepens, van der Slik, & van Hout, 2013). Therefore, we expected weaker co-activation of English in our English-Dutch group than German in our German-Dutch group. As a consequence, we predicted cross-linguistic influence to be weaker or even completely absent in the English-Dutch participants compared to the German-Dutch participants, as in our previous study with bilingual children.

With regard to language overlap, we expected online cross-linguistic influence to be stronger in situations of *partial overlap* than in situations of *complete overlap* and to be absent in situations of *no overlap* (e.g., Hulk & Müller, 2000; van Dijk et al., Chapter 3). Specifically, we expected larger slowdown effects for the V-PP structure in English-Dutch bilinguals and the PP-V structure in the German-Dutch bilinguals compared to the V2 structure in German-Dutch bilinguals. For the V-PP structure, we expected slowdown effects in German-Dutch bilinguals as well, on the assumption that they possess this representation (also see van Dijk et al., Chapter 3). The PP-V, V2 and V3 structures in English-Dutch bilinguals and the V3 structure in German-Dutch bilinguals do not overlap between Dutch and English/German. Therefore, we predicted no cross-linguistic influence for these structures.

With regard to language dominance, we hypothesized that cross-linguistic influence would become stronger the more dominant bilinguals were in the language not in use. In particular, we predicted stronger online effects in our study for more German- and English-dominant participants. This prediction is in line with the online findings in early bilingual adults and the German-Dutch children in our previous study (e.g., Martohardjono et al., 2017; van Dijk et al., Chapter 3).

Finally, with regard to language mode, we expected stronger cross-linguistic influence in a bilingual mode than in a monolingual mode. When participants switch between Dutch and English or German, the latter language becomes more strongly co-activated than when participants are completely in a Dutch monolingual mode (e.g., Grosjean, 1998, 2001; Hopp, 2017) and consequently, cross-linguistic influence should be more apparent.

General bilingualism effect

Our final hypothesis was that if sentence processing in simultaneous bilinguals is less efficient due to them having to control two languages instead of one, as has been claimed for L2 speakers (e.g., Clahsen & Felser, 2006; Polinsky & Scontras, 2020; Sorace, 2011), a general delay should become visible in our bilingual groups compared to monolingually-raised controls. Thus, we expected the English-Dutch and German-Dutch groups to listen to Dutch sentences more slowly than Dutch speakers that grew up monolingually. Crucially, we expected these delays to be similar in the two bilingual groups.

4.3 Method

Participants

We tested 51 simultaneous bilingual adults and adolescents (age of onset of acquisition of both languages before 4;0) who either grew up with English and Dutch (n = 26) or German and Dutch (n = 25). Selection criteria were that at the time of testing participants still received weekly exposure to both of their languages and that (previous) experience with other languages was minimal. More specifically, participants should not have had significant experience with other languages before the age of 4;0; they should not have lived in a country where they had spoken a third language for a longer period than 6 months; and they should not have participated in a bachelor or master program studying a third language. One German-Dutch participant had to be excluded, because she did not receive any exposure in Dutch at the time of testing.

We also tested 31 participants who acquired Dutch monolingually. We will refer to this group as the "Dutch group". From this pool of participants, 25 were chosen who, as a group, matched in age and educational level to the two bilingual groups. As for the bilingual participants, the participants in the Dutch group were also required to have minimal contact with a second language.

Language background questionnaire

All participants were asked to fill out a screening form before testing and a questionnaire after testing. The questionnaire was a mixture of existing questionnaires and adapted such that the variables extracted from it were as comparable as possible as those used in the child study (Bilingual Language Experience Calculator, Unsworth, 2013; Language Experience and Proficiency Calculator, Marian, Blumenfeld, & Kaushanskaya, 2007; Language History Questionnaire, Li, Zhang, Tsai, & Puls, 2014). We collected data about participants' cumulative and current input in their languages, country of birth, and residence and level of education. In case questionnaires were incomplete, we estimated the information based on participants' screening forms when possible.

For our cumulative input measure, we asked participants to estimate the percentage of time they were exposed to their languages for different periods in their life (i.e., during primary school, during secondary school, at university). Using this information, we calculated their cumulative input by adding up their exposure to their languages over time. Current input was based on participants' estimation of the number of hours per week they spent listening to, speaking, reading and writing their languages.

Table 4.3 shows the characteristics of the participant groups. Whilst the groups did not differ significantly in age, they did differ in educational level. Whereas most participants in the Dutch group had finished tertiary education, most English-Dutch and German-Dutch participants had only finished primary or secondary education. Most participants were born in the Netherlands and lived there at time of testing. On average, the onset of acquisition of Dutch was very similar across groups. As expected, past and current input in Dutch was largest in the Dutch group, in English was largest in the English-Dutch group and in German was largest in the German-Dutch group. For current input, similar patterns were found for reading and writing (not included in the table).

Table 4.3 also shows that the distinction between our groups was not entirely categorical (in line with Luk & Bialystok, 2013). First of all, the German-Dutch and Dutch group received exposure to English during childhood. Children growing up in the Netherlands (and Germany) are typically exposed to English from early on in life, for example, through (social) media, songs and playing computer games. This is why some participants in the Dutch and German-Dutch groups indicated they started to listen to English before age 4;0. Furthermore, children are taught English at school in the Netherlands, traditionally starting from the age of 10. Second, most participants in the English-Dutch and Dutch group had some experience with German, because they were taught German at school, from the age of 12, on average. Crucially, however, although not all participants in the Dutch group were truly monolingual during childhood, and participants in all groups were to a certain extent familiar with all three languages, the three groups clearly differed in terms of whether they were functionally bilingual or monolingual in these languages while growing up.

Table 4.3. Overview of background variables for participants (means, standard deviations and ranges).

		English-Dutch	German-Dutch	Dutch	Test statistics
Nr. of participants		26	24	25	
Age at time testing		20.4 (4.9) 15-30	21.9 (7.1) 15-43	23.9 (3.6) 18-34	*F*(2,72) = 2.8
Level of education (n)	Primary	2	4	0	Fisher's exact test***
	Secondary	14	7	10	
	Tertiary	6	8	15	
	Missing	4	5	0	
Country of birth (n)	Netherlands	21	16	25	
	Other	1	7	0	
	Missing	4	1	0	
Country of residence (n)	Netherlands	24	22	25	
	Other	1	2	0	
	Missing	1	0	0	
Age of first exposure (years)	Dutch	0.3 (0.7) 0-2	0.3 (0.8) 0-3	0.04 (0.2) 0-1	*F*(2,72) = 1.6
	English	0.1 (0.4)[a] (0-2)	7.9 (3.7)[b] 0-12	8.6 (2.6)[b] 2-12	*F*(2,72) = 84.1***
	German	12.2 (1.9)[a] 6-14	0.1 (0.4)[b] 0-2	11.3 (2.5)[a] 5-15	*F*(2,57) = 298.8***
Cumulative input (%)	Dutch	57.6 (16.9)[a] 24-80	57.6 (14.4)[a] 29-87	88.1 (5.6)[b] 76-99	*F*(2,65) = 42.4***
	English	41.8 (16.9)[a] 20-76	4.9 (4.8)[b] 0-19	9.7 (4.7)[b] 1-19	*F*(2,65) = 83.2***
	German	0.1 (0.2)[a] 0-1	36.2 (11.5)[b] 9-61	1.0 (0.9)[a] 0-3	*F*(2,65) = 220.0***

Table 4.3 (continued).

		English-Dutch	German-Dutch	Dutch	Test statistics
Current input listening (%)	Dutch	52.7 (20.5)[a] 7-90	60.4 (18.1)[a] 11-88	80.9 (13.5)[b] 49-100	F(2,72) = 17.3***
	English	47.0 (20.6)[a] 10-93	13.3 (12.9)[b] 1-53	18.8 (13.7)[b] 0-51	F(2,72) = 31.6***
	German	0.2 (0.4)[a] 0-1	25.5 (15.6)[b] 3-72	0.1 (0.3)[a] 0-1	F(2,72) = 67.2***
Current input speaking (%)	Dutch	62.8 (24.2)[a] 7-99	67.1 (20.3)[a] 9-97	89.9 (13.5)[b] 49-100	F(2,71) = 13.0***
	English	36.6 (24.6)[a] 1-93	9.1 (13.6)[b] 0-58	9.9 (13.8)[b] 0-51	F(2,71) = 18.5***
	German	0.15 (0.54)[a] 0-2	23.3 (16.2)[a] 3-74	0 (0)[b] 0-0	F(2,71) = 51.5***

Note. ***p < .001; scores with similar subscripts (i.e., a, b) for a certain variable did not differ significantly from one another according to post-hoc tests (Tukey HSD).

Tasks

Self-paced listening task

Three self-paced listening tasks were created to measure participants' online sentence processing (Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996): a completely Dutch task, a task with English and Dutch stimuli and a task with German and Dutch stimuli.

In the Dutch self-paced listening task, experimental items were identical to the stimuli used with bilingual children by van Dijk, Dijkstra and Unsworth (Chapter 3): 15 long passive and 15 adverb-initial sentences in Dutch were recorded in two word orders: PP-V and V-PP, and V2 and V3. For a complete list of stimuli see van Dijk et al. (Chapter 3). Experimental items in the English-Dutch and German-Dutch self-paced listening tasks were based on the stimuli from the Dutch task and were all in Dutch. However, verbs and animal names were used in different combinations to create slightly different sentences. The same experimental items were used in the English-Dutch and German-Dutch task.

Long passives were split into 7 segments and the V2/V3 sentences into 5 segments (see Table 4.4).

Table 4.4. Examples long passive and V2/V3 sentences in self-paced listening task in PP-V, V-PP, V2 and V3 order.

		Segment						
		1	2	3	4	5	6	7
Long passive	PP-V	de leeuw	wordt	door de beer	gekieteld	en	de muis	likt aan een ijsje
		the lion	*is*	*by the bear*	*tickled*	*and*	*the mouse*	*licks an ice cream*
	V-PP	de leeuw	wordt	gekieteld	door de beer	en	de muis	likt aan een ijsje
		the lion	*is*	*tickeld*	*by the bear*	*and*	*the mouse*	*licks an ice cream*
	Comprehension question: Likt de muis?							
	licks the mouse							
	Is the mouse licking?							
Active	V2	op de bank	zingt	de slang	een liedje	van school		
		on the couch	*sings*	*the snake*	*a song*	*from school*		
	V3	*op de bank	de slang	zingt	een liedje	van school		
		on the couch	*the snake*	*sings*	*a song*	*from school*		
	Comprehension question: Is de slang onder de bank?							
	Is the snake under the couch							

The critical region was segments 3 and 4 for the passives and segments 2 and 3 for the V2/V3 sentences, with the spill-over region at segment 5 and segment 4, respectively. All items were recorded by a female native speaker using neutral prosody and intonation and were segmented afterwards. Comprehension questions were asked after 8 passive and 8 V2/V3 items (equal number of yes and no responses); these did not query the critical region itself.

In all three tasks, the 60 experimental items were distributed over pseudorandomized lists, such that every participant heard each item twice: once in the PP-V/V2 order and once in the V-PP/V3 order. To prevent participants from remembering the exact word order of each item when encountering it a second time, each item appeared once in the first half of the experiment and once in the second half.

For the Dutch task, 120 Dutch filler items were created. These consisted of various word orders (72 sentences containing a relative clause; 16 dative constructions; 16 sentences with PP-attachment; and 16 sentences with a particle verb). Twelve fillers were ungrammatical. Filler items were segmented and added to the experimental items from the Dutch experiment. For the English-Dutch and German-Dutch tasks, the Dutch filler items were translated to English and German, respectively, segmented and added to the experimental items.

The experiments were created in E-Prime, version 2.0 (Schneider, Eschmann, & Zuccolotto, 2002). Participants were seated in front of a laptop and button box wearing headphones. At the start of the experiment, they received written instructions informing them that they were going to listen to grammatical and ungrammatical sentences that were cut into fragments. To listen to the entire sentences, they were to press a button in a fast pace. They were also instructed to pay attention to the meaning of the sentences as they would have to answer statements about them. The experiment started with eight practice items. After the practice block the participants were given the opportunity to ask questions. During the experimental block, participants were given the possibility to take a short break after having listened to 60 and 120 items. Experiments took between 20 and 30 minutes to finish.

LexTALE

Participants' proficiency in Dutch, English and German was assessed using the Lexical Test for Advanced Learners of English (LexTALE; Lemhöfer & Broersma, 2012). The LexTALE has been developed as a quick measure of L2 learners' lexical proficiency in Dutch, English and German. For each language participants were shown 40 written words (one-by-one), varying in frequency.

Participants had to indicate whether a word was a real word or a nonsense word. Scores on the task reflected the percentage of items that were answered correct while adjusting for a yes-bias. The written modality of the task could have had a disadvantage in languages in which participants did not receive formal training (at school). However, spelling played little role in the task and the test has been found to correlate well with lexical and general language proficiency in L2 learners (e.g., Lemhöfer & Broersma, 2012). Therefore, we considered the task suitable for our population.

Digit span

Participants' verbal short-term and working memory abilities were assessed using a forward and backward digit span task in Dutch (Automated Working Memory Assessment (AWMA): Alloway, 2012). The standard scoring procedure of the AWMA was used (forward: max. 48; backward: max. 36).

Procedure

Bilingual participants were tested during two sessions at the university or at home. The first session was in a completely Dutch monolingual mode. Participants started with the Dutch self-paced listening task, which was followed by the digit span and Dutch LexTALE task. The second session was in a bilingual mode: English-Dutch for the English-Dutch group and German-Dutch for the German-Dutch group. The session started with the English-Dutch or German-Dutch self-paced listening task. This was followed by the English and German LexTALE tasks (the English task came first in the English-Dutch group and the German task came first in the German-Dutch group). The Dutch group was tested during one test session. The English and German LexTALEs directly followed the Dutch one.

Participants filled out the questionnaires by themselves at a different moment. All participants gave written consent and were rewarded a voucher of €10,- per test session.

Data preparation

The data from the long passives and V2/V3 sentences were analysed separately. The critical segments differed in audio length between conditions due to the difference in word orders. This is a common issue in self-paced listening and reading experiments (e.g., Chondrogianni, Marinis, Edwards, & Blom, 2015; Ferreira & Clifton, 1986). Following standard procedures (e.g., Marinis, 2010), we therefore calculated residual reaction times (RTs) by subtracting the duration of each audio fragment from participants' raw RTs.

Extreme residual RTs above 2000 ms as well as segments that had been listened to less than 300 ms were removed. Residual RTs were log-transformed to correct for the positively skewed distribution of RTs. Next, average residual RTs of participants and items were inspected. No participant or item deviated more than 2.5 SDs from the group/item means and no participant or item was removed from the dataset. Finally, residual RTs that deviated more than 2.5 SDs from individual participants' average residual RTs by condition by segment were removed. In total, less than 5% of the residual RTs were removed.

A relative proficiency score for participants was calculated by subtracting English-Dutch participants' LexTALE score in English from their LexTALE score in Dutch and by subtracting German-Dutch participants' LexTALE score in German from their LexTALE score in Dutch. A positive proficiency score thus indicated higher proficiency in Dutch, a negative score higher proficiency in English/German and scores around zero indicated balanced proficiency across languages. For comparison, a difference score was calculated for the Dutch group as well (Dutch LexTALE score – English LexTALE score).

4.4 Results

Background tasks

Table 4.5 shows participants' performance on the LexTALE and digit span tasks. Scores on the LexTALE reflect the patterns found for participants' current and cumulative input to their languages. To be more precise, proficiency in Dutch as measured by the LexTALE was significantly higher in the Dutch group compared to the two other groups. Similar results were found for English in the English-Dutch group and German in the German-Dutch group. The bilingual groups' Dutch LexTALE scores were comparable, however. Furthermore, the English-Dutch participants' English LexTALE scores were very similar to the German-Dutch participants' German LexTALE scores. The average relative proficiency of the bilingual groups in their two languages was close to zero showing relatively balanced proficiency in both languages on average. The Dutch group, in contrast, was on average more proficient in Dutch than in English.

There were no significant differences in participants' digit span scores. This indicates that regardless of differences in educational level the groups were comparable in terms of short-term and working memory capacity.

Table 4.5. Participants' average accuracy scores, standard deviations and ranges on the LexTALE and Digit span tasks.

		English-Dutch	German-Dutch	Dutch	ANOVA
LexTALE (%)	Dutch	78.6 (8.5)[a] 57.8-89.7	82.6 (11.0)[a] 56.0-96.4	88.0 (6.0)[b] 78.3-100.0	F(2,72) = 7.5**
	English	82.9 (11.7)[a] 53.2-100	67.7 (14.1)[b] 39.2-96.1	70.7 (15.3)[b] 49-96.4	F(2,71) = 8.4***
	German	53.2 (9.9)[a] 34.1-71.0	77.3 (11.5)[b] 56.6-100	55.5 (9.8)[a] 39.2-74.5	F(2,71) = 39.7***
	Relative proficiency	-4.5 (10.6)[a] -22.0-12.8	5.3 (11.3)[b] -33.0-24.7	17.3 (11.8)[c] -3.6-36.5	F(2,71) = 23.6***
Digit span	Forward	33.3 (7.0) 24-48	32.9 (5.6) 25-48	34.2 (5.7) 24-46	F(2,72) = 0.3
	Backward	21.5 (6.1) 13-32	21.9 (5.1) 14-34	20.5 (5.9) 10-33	F(2,72) = 0.4

Note. ***p* < .01; ****p* < .001; scores with similar subscripts (i.e., a, b, c) for a certain variable did not differ significantly from one another according to post-hoc tests (Tukey HSD).

Self-paced listening task

All self-paced listening data were analysed using multi-level linear effects models in R (version 4.0.3, R Core Team, 2020; package lme4, version 1.1-23, Bates, Maechler, Bolker, & Walker, 2015; and lmerTest, version 3.1-2, Kuznetsova, Brockhoff, & Christensen, 2017). An alpha level of .05 was used for all statistical tests. Separate analyses were run for the long passives and V2/V3. All models contained random slopes by Word order and Segment for individual participants and random intercepts for participants and items. In case of singularity or convergence issues random slopes and intercepts were dropped from the model until issues were resolved. For all models reported, residuals exceeding 2.5 had been removed using the LMERConvenienceFunctions package (version 3.0, Tremblay & Ransijn, 2020; < 4% of data removed).

In a first step, base models were created for residual RTs on the pre-critical segment, the critical segments and the spill-over segment of the long passive (segment 2, 3, 4 and 5) and V2/V3 conditions (segment 1, 2, 3 and 4). All models contained grand mean centred fixed effects of *RT on the previous trial*, *Duration of the audio fragment* and *Trial number*. Age of the participants did not significantly predict residual RTs and was therefore dropped. *Digit span forward* only significantly predicted participants residual RTs in session 1 and was therefore dropped as fixed effect from the models for session 2. The interaction between Segment and Word order was added to all models.

In a second step, we tested our hypotheses. First, we added the fixed effect of Group to our base models, to test whether the bilingual groups performed differently from the Dutch group. Second, we tested whether the 3-way interaction between Group, Segment and Word order was significant, indicative of cross-linguistic influence. A main effect or interaction was deemed significant when it significantly improved a model without the effect or interaction based on log likelihood tests. Helmert contrasts were used to explore significant effects and interactions. Specifically, model summaries first compared residual RTs of the bilingual groups to the Dutch group to test for a general effect of bilingualism. Residual RTs of the English-Dutch group were then compared to the residual RTs of the German-Dutch group to test for cross-linguistic influence. Models were relevelled when necessary.

In a final step, we tested for effects of language dominance in the two bilingual groups separately at the critical and spill-over segments. Proxies of language dominance were percentage *Current input* (listening), *Current output* (speaking) and percentage *Cumulative input* to English in the English-Dutch group and German in the German-Dutch group as well as participants' *Relative proficiency* in their languages. These predictors were tested in separate analyses: (1) as a fixed effect; (2) in interaction with *Word order*; and (3) in interaction with *Word order* and *Group*. Again, significance of main effects and interactions were tested by means of log likelihood tests and further explored by model summaries.

Monolingual session

Long passives - main analyses. Figure 4.1 shows participants' residual RTs in the PP-V and V-PP condition (for the average residual RTs and standard deviations, see A4.1 in the appendix). Because the groups showed slightly different patterns in the PP-V and V-PP word orders at the two pre-critical segments 1 and 2, we decided to average residual RTs from both segments and use those in the analyses. This way behaviour across groups at the pre-critical region was directly comparable.

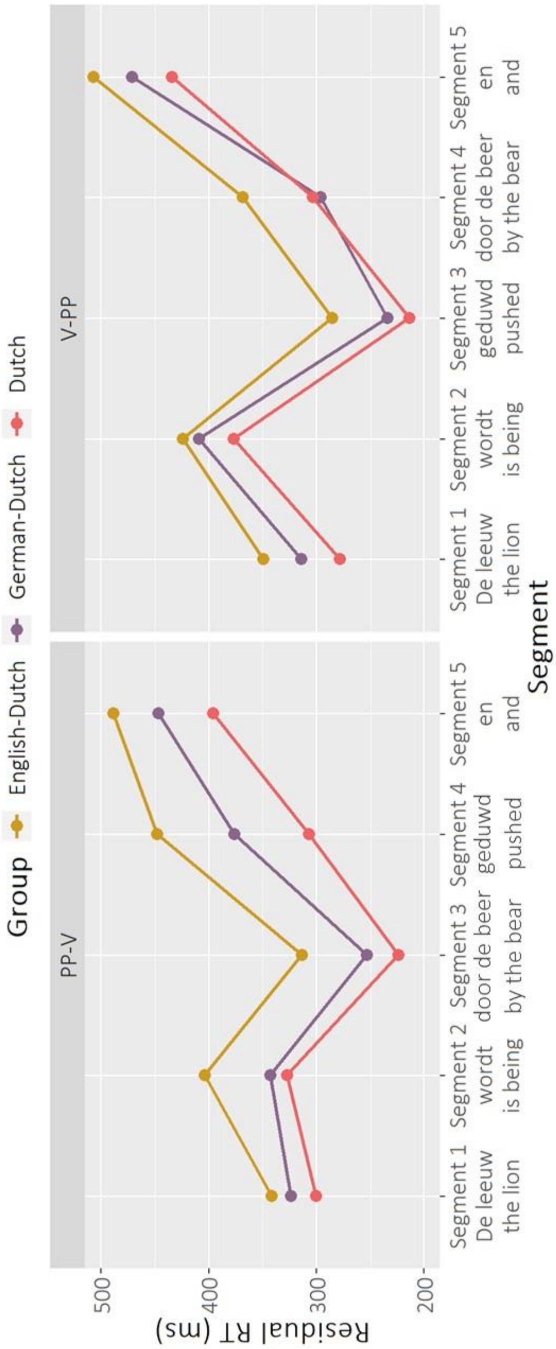


Figure 4.1. Average residual RTs in the PP-V (left panel) and V-PP (right panel) condition by *Segment* by *Group* in the monolingual session.

Overall, the Dutch group had slightly smaller residual RTs than the bilingual groups. Furthermore, the English-Dutch group was slightly slower than the German-Dutch group. The main effect of *Group* was not significant, however (X^2 = 2.4; *Δdf* = 2; *p* = .305).

The 3-way interaction between *Group*, *Word order* and *Segment* was significant (X^2 = 15.8; *Δdf* = 6; *p* = .015). Table 4.6 shows summary effects of *Word order* in interaction with *Group* at each segment. The effect of *Word order* was similar in the group comparisons at pre-critical segments 1 and 2, critical segment 3 and spill-over segment 5. However, at critical segment 4 the effect of *Word order* was significantly different in the two bilingual groups compared to the monolingual group. The bilingual groups did not differ significantly from each other, however.

Table 4.6. Simple interactions between *Group* and *Word order* in the long passive condition at segments 2, 3, 4 and 5. The model was relevelled based on *Group* and *Segment*.

		B	*SE*	*t*	*p*
Segment 1&2	Bi- vs. monolingual	-0.0042	0.0052	-0.8	.423
de leeuw wordt	English vs. German	0.0028	0.0061	0.5	.647
(the lion is being)					
Segment 3	Bi- vs. monolingual	-0.0044	0.0052	-0.9	.396
door de beer/geduwd	English vs. German	-0.0053	0.0061	-0.9	.387
(by the bear/pushed)					
Segment 4	Bi- vs. monolingual	0.0161	0.0053	3.1	.002
geduwd/door de beer *(pushed/by the bear)*	English vs. German	-0.0037	0.0062	-0.6	.548
Segment 5	Bi- vs. monolingual	0.0007	0.0052	0.1	.895
en *(and)*	English vs. German	-0.0065	0.0060	-1.1	.280

In order to test for a general effect of bilingualism, we explored the significant interaction at segment 4 by comparing the bilingual groups' results in the PP-V and V-PP conditions separately to the results of the monolingual group. Summary effects showed that the bilingual groups' residual RTs did not differ significantly from the residual RTs of the Dutch group at segment 4 in the PP-V (*B* = -0.022; *SE* = 0.013; *t* = -1.632; *p* = .107) or V-PP condition (*B* = -

006; *SE* = -0.013; *t* = -0.418; *p* = .677; see A4.2 in the appendix for the complete model summaries). In other words, whilst there was a numerical difference at segment 4 in the PP-V condition between the bilingual groups and the monolingual group and in the V-PP condition between the English-Dutch group and the monolingual group, these differences did not reach significance in the analyses. This was probably due to the large standard deviations within the groups.

Long passives – Language dominance. Language dominance did not significantly improve any model (see A4.3 in the appendix for log likelihood tests for the different models).

Verb second – main analyses. Figure 4.2 shows participants' residual RTs in the V2 and V3 condition (for the average residual RTs and standard deviations, see A4.1 in the appendix). Similar to the long passive sentences, the bilingual participants had slightly larger residual RTs than the monolingual group. Furthermore, the English-Dutch participants were slightly slower than the German-Dutch participants. Again, however, the main effect of *Group* was not significant (X^2 = 1.5; *Δdf* = 2; *p* = .466). The 3-way interaction between *Group*, *Word order* and *Segment* was not significant either (X^2 = 6.5; *Δdf* = 6; *p* = .370). This shows that the groups' behaviour did not differ significantly from each other.

Verb second – language dominance. *Current listening, Cumulative input* and *Relative proficiency* did not significantly improve any model in neither group (see A4.3 in the appendix for log likelihood tests for the different models). In contrast, *Current speaking* did interact significantly with *Word order* in both the English-Dutch (X^2 = 5.9; *Δdf* = 1; *p* = .016) and German-Dutch group (X^2 = 4.1; *Δdf* = 1; *p* = .042).

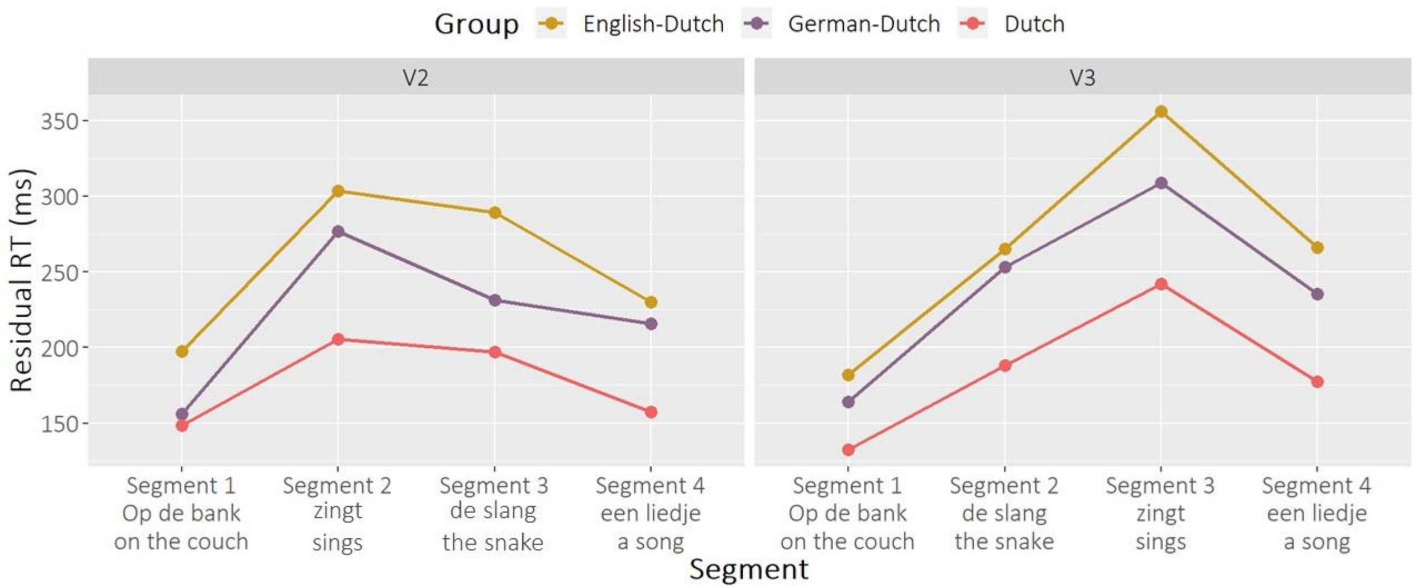


Figure 4.2. Average residual RTs in the V2 (left panel) and V3 (right panel) condition by *Segment* by *Group* in the monolingual session.

Figure 4.3 shows the estimated marginal means of the interaction between *Current speaking* and *Word order* in the English-Dutch group. When averaged over segments, the English-Dutch adults were slightly faster in the V2 condition compared to the V3 condition. However, this difference between conditions became smaller, the more English relative to Dutch participants were speaking at the time of testing. The effect of *Current speaking* differed significantly between conditions (B = -0.0004; SE = 0.0002; t = -2.450; p = .022), but simple effects of *Current speaking* in the V2 and V3 word order were not significant (V2: B = 0.0000; SE = 0.0005; t = 0.135; p = .894; V3: B = -0.0003; SE = 0.0005; t = -0.597; p = .556). Hence, whilst the effect of *Current speaking* was significantly different between conditions, it failed to significantly predict English-Dutch participants' listening times in the two conditions separately.

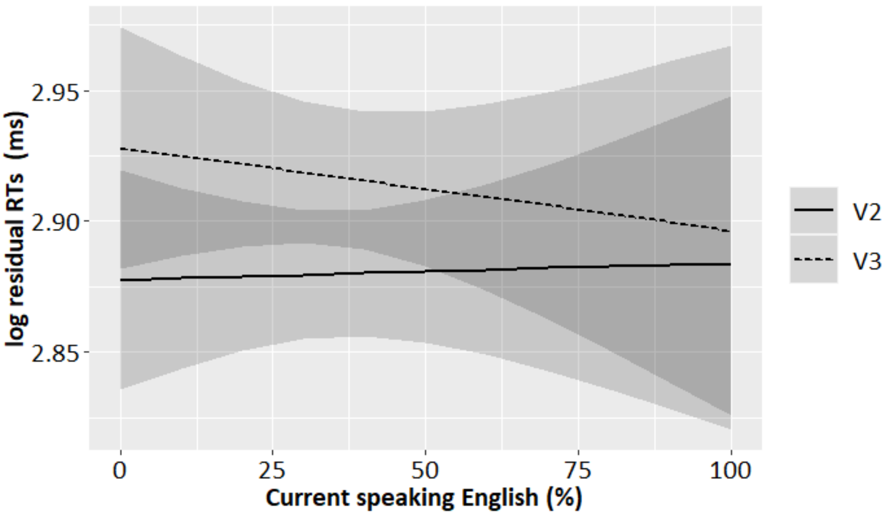


Figure 4.3. Estimated marginal means of English-Dutch participants' residual RTs in the V2 and V3 sentences by *Current speaking* of English averaged over segments 2, 3 and 4.

Figure 4.4 shows the estimated marginal means of the interaction between *Current speaking* and *Word order* in the German-Dutch group. Similar to the English-Dutch group, the German-Dutch adults were overall slightly faster in the V2 condition compared to the V3 condition. However, in contrast to the English-Dutch group, this difference became *larger* the more German participants spoke at the time of testing. Summaries showed that the more German spoken, the smaller the residual RTs became in the V2 condition (B = -0.0002; SE = 0.0012; t = -0.151; p = .882) and the larger in the V3 condition (*B* = 0.0006; *SE* = 0.0011; *t* = 0.564; *p* = .579). These effects were not significant, however. Furthermore, the interaction between *Current speaking* and *Word order* did not reach significance (*B* = 0.0008; *SE* = 0.0004; *t* = 2.027; *p* = .057). In short, *Current speaking* did not significantly predict German-Dutch participants' listening times.

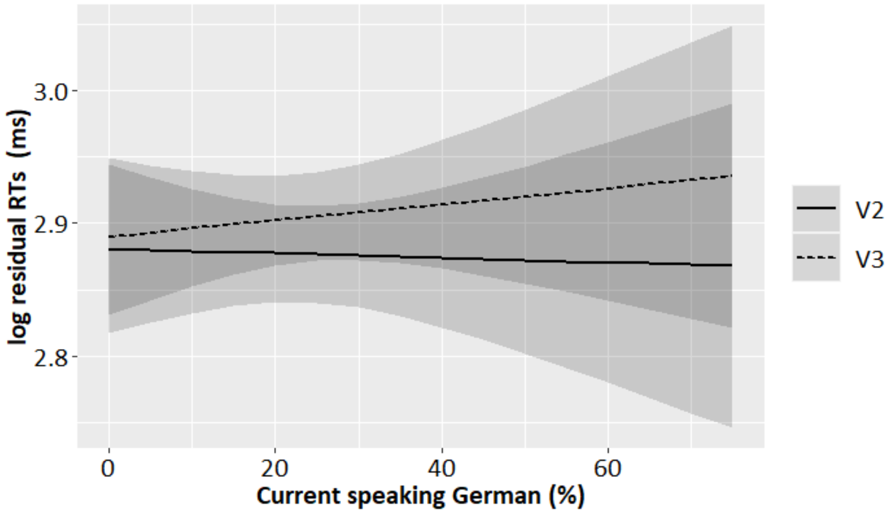


Figure 4.4. Estimated marginal means of German-Dutch participants' residual RTs in the V2 and V3 sentences by *Current speaking* of German averaged over segments 2, 3 and 4.

Bilingual session

Long passives – main analyses. Figure 4.5 shows the English-Dutch and German-Dutch participants' residual RTs in the PP-V and V-PP condition from session 2 (for the average residual RTs and standard deviations, see A4.1 in the appendix). Similar to session 1, the English-Dutch participants were slightly slower than the German-Dutch participants. However, the main effect of *Group* was not significant (X^2 = 1.0; *Δdf* = 1; *p* = .323). Furthermore, the 3-way interaction between *Group*, *Condition* and *Segment* was not significant either (X^2 = 2.9; *Δdf* = 3; *p* = .412). This indicates that the listening times of the two bilingual groups on the long passive sentences did not differ significantly in either word order.

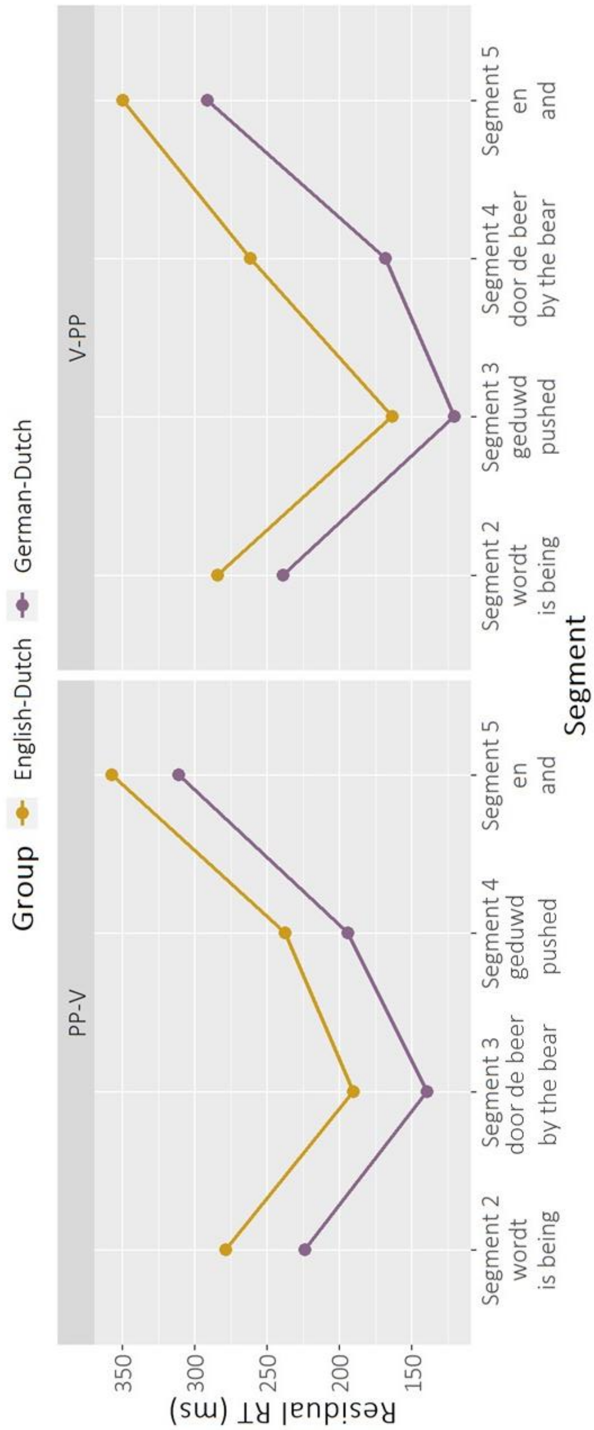


Figure 4.5. Average residual RTs in the PP-V (left panel) and the V-PP (right panel) condition by *Segment* by *Group* in session 2.

Long passives – language dominance. In the English-Dutch group, there were no significant effects of or interactions with our dominance measures (see A4.3 in the appendix for log likelihood tests for the different models). In the German-Dutch group the 3-way interaction between *Cumulative input*, *Word order* and *Segment* was significant (X^2 = 15.9; Δdf = 2; p < .001). None of the other interactions with our dominance measures were significant.

Figure 4.6 shows the estimated marginal means of the interaction between *Cumulative input, Word order* and *Segment* in the German-Dutch group. At segment 3 and 5, the interaction between *Cumulative input* and *Word order* was not significant (segment 3: B = 0.0006; SE = 0.0004; t = 1.676; p = .097; segment 5: B = 0.0006; SE = 0.0001; t = 0.288; p = .774). At segment 4, the interaction between *Cumulative input* and *Word order* was significant (B = 0.0012; SE = .0004; t = -3.396; p < .001). In both conditions German-Dutch bilinguals became slower the more exposure they had received to German in the past. However, this effect was significant only in the PP-V condition (PP-V: B = 0.0002; SE = 0.0008; t = 2.665; p = .014; V-PP: B = 0.0008; SE = 0.0007; t = 1.056; p = .302). In sum, *Cumulative input* significantly predicted German-Dutch participants' listening times in the PP-V word order at critical segment 4.

V2 – main analyses. Figure 4.7 shows the English-Dutch and German-Dutch participants' residual RTs in the PP-V and V-PP condition from session 2 (for the average residual RTs and standard deviations, see A4.1 in the appendix). Although the English-Dutch participants had overall larger residual RTs than the German-Dutch participants, the main effect of *Group* was not significant (X^2 = 0.6; Δdf = 1; p = .424). The 3-way interaction between *Group*, *Word order* and *Segment* was not significant either (X^2 = 3.6; Δdf = 3; p = .314). This indicates that the two bilingual groups had similar listening patterns in both word orders.

V2 – language dominance. None of the effects or interactions reached significance (see A4.3 in the appendix for log likelihood tests for the different models).

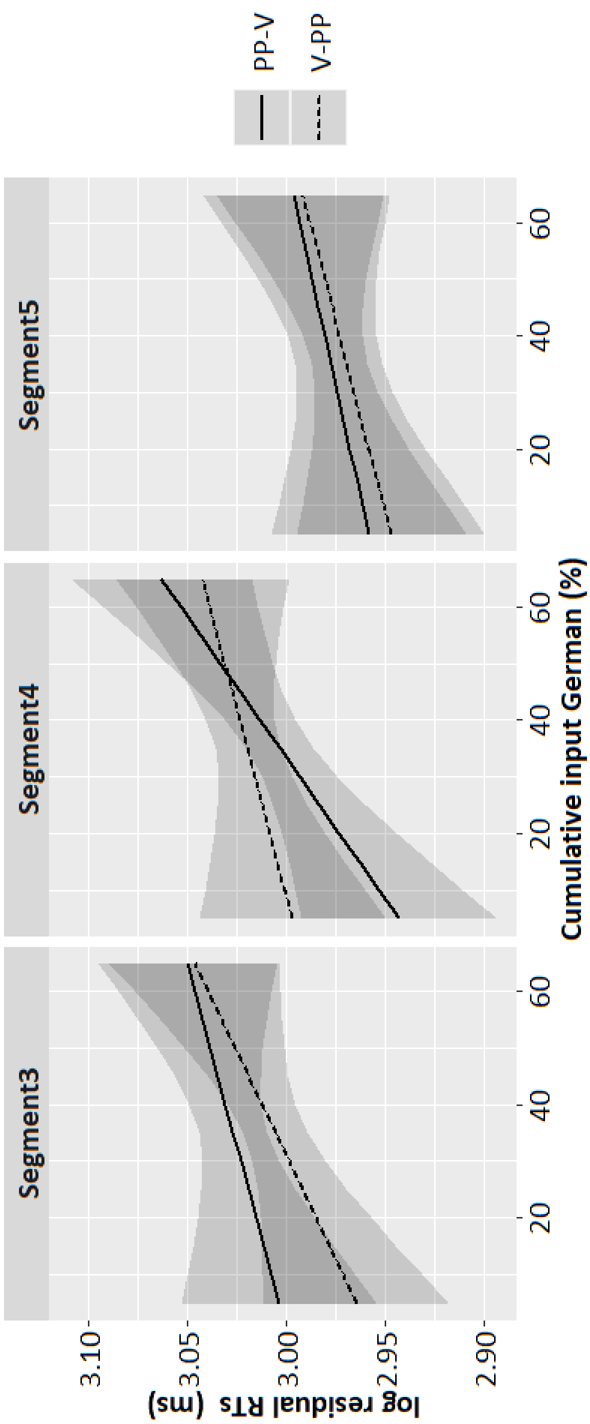


Figure 4.6. Estimated marginal means of German-Dutch participants' residual RTs in the PP-V and V-PP sentences by *Cumulative input* in German at segments 3, 4 and 5.

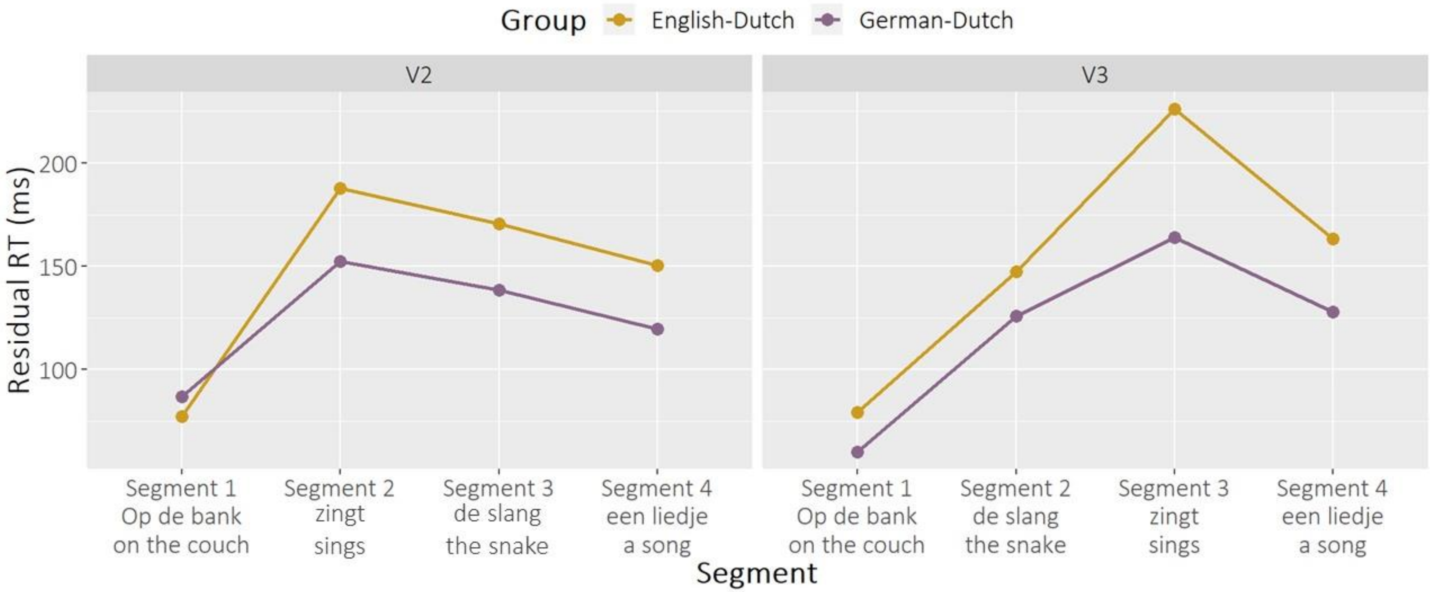


Figure 4.7. Average residual RTs in the V2 (left panel) and V3 (right panel) condition by *Segment* by *Group* in session 2.

4.5 Discussion

By means of a self-paced listening task we investigated cross-linguistic influence during sentence processing in simultaneous bilingual adults and adolescents. We found that bilingual participants slowed down when listening to sentences in Dutch that overlapped in word order with their other language. This finding supports our first hypothesis that cross-linguistic influence leads to online slowdown effects. Furthermore, the observed effect of online-crosslinguistic influence was moderated by lexical overlap, surface overlap, language dominance and language mode. In particular, slowdown effects were only observed in the German-Dutch participants, in the partially overlapping PP-V word order, in relationship to language dominance and in the bilingual test session. These observations corroborated our second hypothesis that our four predictors of cross-linguistic influence (i.e., lexical overlap, surface overlap, language dominance and language mode) moderate the strength of the slowdown effect. Finally, whilst the bilingual groups were numerically slower than the Dutch group, this difference between groups did not reach significance in any of the analyses. Hence, we found no evidence for our third and final hypothesis that being bilingual results in slower sentence processing.

In the remainder of this section, we discuss our findings in relation to our previous study with bilingual children (van Dijk et al., Chapter 3) and other literature. Furthermore, we explain our observations in terms of language co-activation and inhibition. We show how such an account can explain, first, the similarities in online outcomes between this study and the child study, and, second, the differences between these studies. Throughout this discussion, we refer to the adults and adolescent as adults.

Similarities between bilingual adults and children

Online cross-linguistic influence and language co-activation

Online cross-linguistic influence occurred in the same direction in the bilingual adults and children in our studies, namely as a slowdown effect during listening. This finding suggests that similar mechanisms underlying cross-linguistic influence are at play in different bilingual populations. Following Hopp (2017), we believe that these mechanisms are language co-activation and inhibition. In particular, our adult and child results suggest that during sentence processing in one language, overlapping word orders in the other language become co-activated. As a consequence, the listener has to allocate processing resources to inhibit this co-activation. In turn, fewer processing resources are temporally available for sentence parsing. As a result, the processing of a sentence can become delayed, which is reflected by the slowdown effects during listening in our studies.

Our online findings also fit with those in other online studies with early bilingual adults (e.g., Martohardjono et al., 2017; Runnqvist et al., 2013) and adult L2 learners (e.g., Foucart & Frenck-Mestre, 2012; Hopp, 2017) and in cross-language priming studies with adults and children (e.g., Hartsuiker & Bernolet, 2017; Hsin, Legendre, & Omaki, 2013; Vasilyeva et al., 2010). In particular, these studies all suggest that a structure in a bilingual's one language can activate a similar structure in their other language. In production tasks, such co-activation may facilitate the production of an overlapping sentence structure (e.g., Hartsuiker & Bernolet, 2015; Runnqvist et al., 2013; Vasilyeva et al., 2010). In comprehension tasks, such co-activation may result in less efficient sentence processing (e.g., Foucart & Frenck-Mestre, 2012; Hopp, 2017; Martohardjono et al., 2017). Importantly, structural co-activation seems to underlie effects in different tasks in different populations.

Predictors of cross-linguistic influence

The same variables that predicted the presence and strength of online cross-linguistic influence in simultaneous bilingual children applied to simultaneous bilingual adults, namely, lexical overlap, surface overlap and language dominance. This observation further supports our proposal that the same mechanisms underlie cross-linguistic influence in bilingual adults and children. In particular, we propose that each of the three variables contributes to the extent of language co-activation during sentence processing and, in turn, the extent to which inhibition is necessary to suppress this co-activation. Furthermore, language mode, not tested in the child study, predicted cross-linguistic influence in the bilingual adults. We will now discuss each predictor in relation to co-activation and inhibition.

Lexical overlap. In our view, the more lexical overlap bilinguals' languages share, the stronger the language not in use becomes activated during the processing of another language (also see van Dijk et al., Chapter 3). This is in line with lexical boost effects observed in structural priming studies (e.g., Pickering & Branigan, 1998). With regard to our participant groups, German shares more lexical overlap with Dutch than English with Dutch (e.g., Schepens et al., 2013). As a consequence, during the self-paced listening task the overall co-activation of German should have been higher than of English. Therefore, more processing resources had to be allocated to inhibit co-activation of German than of English. The results for the bilingual adults and children suggest that only the co-activation of German was large enough to cause visible inhibition effects during sentence processing. In fact, although we expected that sentence processing in Dutch co-activated overlapping structures in English and German, only in German was this co-activation apparently strong enough to result in visible slowdown effects in our self-paced listening task.

There is some evidence from studies with child and adult L2 learners that (the absence of) lexical overlap in experimental items predicts online cross-linguistic influence (e.g., Hopp, 2017; Lemmerth & Hopp, 2019). Future studies with simultaneous bilinguals should manipulate the amount of lexical overlap between sentences to investigate the exact role of lexical overlap during simultaneous bilingual sentence processing.

Surface overlap. We consider the type of surface overlap a proxy for the relative frequency with which a particular structure occurs in bilinguals' languages. In turn, this relative frequency should predict the level of co-activation of a structure during sentence processing (van Dijk et al., Chapter 3; also see Runnqvist et al., 2013). In the case of *complete overlap*, we assume that a structure is equally frequent in both languages of a bilingual listener. In the case of *partial overlap*, we assume that a structure is more frequent in the language of a bilingual listener with only one option than in the language with two options. In the case of *no overlap*, a structure is only present in one language of a bilingual listener and not in the other. Hence, the relative frequency of a structure is not relevant in the latter situation. In contrast, as a consequence of their frequency distributions, we expect language co-activation to be stronger in situations of *partial overlap* relative to *complete overlap*. In our view, the processing of the Dutch PP-V structure co-activates the German PP-V structure to a larger extent than the processing of the Dutch V2 structure co-activates the German V2 structure. As a result, more processing resources need to be allocated to inhibit partially overlapping structures than completely overlapping structures. Hence, cross-linguistic influence is more likely to occur online in a situation of *partial overlap* than in a situation of *complete overlap*. Furthermore, non-overlapping structures, such as the ungrammatical V3 structure in Dutch, should not be able to co-activate structures in bilinguals' other language. Therefore, cross-linguistic influence should not occur in a situation of no *overlap*.

Differences in co-activation and inhibition between overlap situations thus account for the more pronounced effects of online cross-linguistic influence in *partial overlap* situations than in *complete overlap* situations for German-Dutch children, and for the presence of online cross-linguistic influence only in *partial overlap* situations for German-Dutch adults. More specifically, in the children co-activation of the partially overlapping PP-V and V-PP structures resulted in slowdown effects visible irrespective of children's dominance profiles. In contrast, co-activation of the completely overlapping V2 structure in German was only strong enough to become visible when the children's dominance profile was taken into account (discussed in more detail in the next section). In German-Dutch adults, only co-activation of the partially overlapping PP-V structure in German, and not the completely overlapping V2 structure, was strong enough to become visible in the self-paced listening task. Furthermore, the absence of co-activation in *no overlap* situations explains why we did not observe online cross-linguistic influence for non-overlapping structures in German-Dutch children and adults.

Our observations for surface overlap are in line with offline and online studies that attested cross-linguistic influence with *partial overlap* in early bilingual adults (e.g., Anderssen et al., 2018; Kupisch, 2014; Runnqvist et al., 2013) and in simultaneous bilingual children (e.g., Haznedar, 2007; Hulk & Müller, 2000; Meroni et al., 2017). However, our findings contrast with effects of cross-linguistic influence in situations of *no overlap* in online studies with early bilingual adults (e.g., Kupisch & Barton, 2013; Martohardjono et al., 2017). For example, Martohardjono and colleagues (2017) found that with increased exposure to English, bilingual participants became less sensitive to ungrammatical *that*-trace omissions in Spanish, as measured by an ERP task. Whilst in English, *that*-trace omissions are allowed, *that*-traces are obligatory in Spanish. Hence, based on our account of co-activation, we would predict that ungrammatical Spanish sentences without a *that*-trace should not be able to activate the grammatical structure in English. However, the findings by Martohardjono and colleagues suggest that co-activation of non-shared structures is possible. This is also in line with observations that bilingual children sometimes use or accept ungrammatical structures from their one language into their other, such as V3 orders in German or V2 orders in English (e.g., Bosch & Unsworth, 2020; Döpke, 1998). Possibly, in our study, the ungrammatical V3 order in Dutch did activate this order in English in English-Dutch participants. Nevertheless, this co-activation may not have been strong enough to affect participants' listening times, as for the other words orders in the English-Dutch group. Potentially, the ERP technique Martohardjono and colleagues (2017) used might be better suited to pick up such subtle effects of co-activation. Future studies should therefore further compare effects of cross-linguistic influence in different overlap situations during sentence processing in simultaneous bilingual adults using different online techniques.

Language dominance. We believe that the more dominant bilingual adults and children are in the language not in use, the stronger this language is co-activated during sentence processing. In turn, the more the language not in use is co-activated, the more processing resources have to be allocated for inhibition. Therefore, with increased dominance, cross-linguistic influence is more likely to become stronger and visible during sentence processing. Hence, co-activation explains why slowdown effects in our self-paced listening task became stronger in the PP-V structures the more dominant adults were in German in the same way we found that these effects became stronger in the PP-V and V2 structures the more dominant children in our earlier study were in German.

Our results for language dominance are in line with studies with early bilingual adults that observed cross-linguistic influence only in the non-dominant language (e.g., Kupisch, 2012; Montrul, 2010) or that observed stronger cross-linguistic influence with increasing dominance in the language not in use (e.g., Martohardjono et al., 2017). Furthermore, our results are also in line with online studies with adult L2 learners (e.g., Foucart & Frenck-Mestre, 2012; Hopp, 2017) and with offline comprehension and production studies with bilingual children (e.g., van Dijk et al., Chapter 2) that found language dominance to affect the presence and strength of cross-linguistic influence.

Language mode. Finally, language mode further predicted effects of cross-linguistic influence in the bilingual adults. Again, this can be accounted for by language co-activation. In our bilingual test session, the German-Dutch adults constantly had to switch between Dutch and German. As a consequence, German should have been more highly activated in the bilingual session than in the monolingual Dutch session (e.g., Hopp, 2017). This explains why we observed online cross-linguistic influence in the bilingual adults only in the bilingual session and not in the monolingual session: only in the bilingual session the co-activation of German was large enough to result in slowdown effects during listening. Our observations are in line with Hopp (2017) who observed cross-linguistic influence during reading in highly proficient L2 learners in a bilingual but not in a monolingual language mode.

General bilingualism effect

There was no evidence for a general bilingualism effect resulting in delays during the self-paced listening task for bilingual adults. This was also the case for the children. In other words, we did not observe any significant delays in the bilingual groups compared to the Dutch control groups other than the ones we attributed to effects of cross-linguistic influence. Our findings therefore show that proposals about less efficient processing in adult L2 learners cannot automatically be extended to simultaneous bilinguals (e.g., Polinsky & Scontras, 2020; Sorace, 2011), at least in situations in which simultaneous bilinguals use their majority language, as was the case here. Future research is necessary to investigate bilingualism effects in simultaneous bilinguals' minority language (also see Felser, 2020).

Differences between bilingual adults and children

In addition to the aforementioned similarities, we also observed two differences between the German-Dutch bilingual adults and the German-Dutch bilingual children. First, cross-linguistic influence was less pronounced in the adults than in the children. Second, different measures of language dominance predicted cross-linguistic influence in adults and children. In this section we discuss these differences in more detail.

Online cross-linguistic influence in bilingual adults and children

Whilst cross-linguistic influence was found for only one structure in the adult bilinguals (i.e., the PP-V structure), it was demonstrated for three structures in the child bilinguals (i.e., the PP-V, V-PP and V2 structures). Furthermore, online cross-linguistic influence was only attested in the adult bilinguals when their language dominance profile was taken into account. In contrast, online cross-linguistic influence in the PP-V and V-PP structures in the bilingual children was found regardless of language dominance. We entertain two complementary explanations for the more pronounced effects of cross-linguistic influence in the bilingual children.

First of all, the bilingual adults likely had more processing resources available for inhibition than the children by virtue of their more advanced age (see Schneider, 2015 for a review of literature on the development of working memory in children). This means that adults may typically have sufficient processing resources available for sentence processing and inhibiting co-activation in parallel. Consequently, co-activation and inhibition were less likely to result in delays during listening in the adults than the children. Furthermore, the adults had more years of experience with processing their languages. Therefore, they will likely have been more trained in inhibiting their other language than the children. Consequently, inhibition may have been more efficient in the adult group, again resulting in smaller and fewer delays. There is indeed some evidence for the importance of language experience over time in adult bilinguals: only cumulative input significantly predicted the German-Dutch adults listening times in the PP-V condition, and not the other dominance measures.

Importantly, our account explains why offline cross-linguistic influence is more likely to surface as overcorrection in bilingual adults than in children. Because adults have sufficient resources to inhibit the language not in use, they might completely suppress offline responses that converge with the language not in use. For instance, in Kupisch's (2014) study, Italian-German bilinguals were found to correct grammatical prenominal adjective-noun orders in Italian. On our account, this is because the prenominal adjective-noun order is the canonical order in German and is therefore inhibited in Italian (also see Anderssen et al., 2018 for a similar account). In contrast, children may not always have sufficient resources to inhibit co-activation, resulting in offline responses that converge with the language not in use. This could explain, for instance, the overuse of prenominal adjective-noun orders in French under influence of English (Nicoladis, 2006).

Second, most of the German-Dutch participants we tested were exposed to and used Dutch on a day-to-day basis more frequently than German. In the children there was a larger range in dominance patterns with most children receiving extensive daily input in German at home. Such differences in exposure patterns between younger and older bilinguals are typical for heritage speakers, who learn a minority language from their parents in a majority language context outside the home (e.g., Montrul, 2010; Polinsky & Kagan, 2007). This implies that German will likely have been more active in the bilingual children tested in our previous study than in the bilingual adults in this study. It also explains why effects of co-activation only showed up in the German-Dutch bilinguals when co-activation of German was maximized: in a situation of *partial overlap*, with increased dominance in German, and in a bilingual language mode.

In sum, cross-linguistic influence was less pronounced in the adults in the current study than in the children in our previous study. We believe that this is the result of quantitative differences between bilingual adults and children in terms of their processing capacity, the years of experience with their languages, and the specific dominance profiles present in our two samples. We believe that the mechanisms behind online cross-linguistic influence are, however, qualitatively similar between the two groups.

Measures of language dominance

Language dominance predicted the strength of cross-linguistic influence in the PP-V condition in both the adults and the children but the measure used to operationalise language dominance differed. More specifically, for adults cumulative input predicted the strength of the effect and for children this was relative proficiency. The question is why different measures predicted cross-linguistic influence in adults and children. We entertain a number of explanations.

First, it is important to note that in the child group a similar though non-significant trend was observed for cumulative input (p = .059) as we found for children's relative proficiency. Hence, this suggests that cumulative input predicts online cross-linguistic influence in both adults and children, but to a lesser extent in the latter group. However, the observed difference between groups may also be a consequence of different patterns of variance in the cumulative input measure. These various options need to be disentangled.

Second, we used different measures of proficiency for adults and children. For adults, we used a written lexical task, whereas for children we used a spoken sentence repetition task. Consequently, our adult proficiency measure was lexical in nature and may have drawn on written language knowledge (but see Lemhöfer & Broersma, 2012). In contrast, our child proficiency measure also included a syntactic component (e.g., Polišenská, Chiat, & Roy, 2015) and did not involve written language knowledge. The latter measure might have been a more direct predictor of bilinguals' online behaviour, given that we measured cross-linguistic influence at the syntactic level using a spoken task. Unfortunately, we were not aware of any (short) syntactic proficiency task suitable for adults and available in our languages of interest that we could include in our test battery. Additional research with simultaneous bilingual adults is therefore necessary including a more syntactic proficiency measure.

Third, it is possible that in adult simultaneous bilinguals, relative proficiency is no longer relevant for effects of cross-linguistic influence. In the child study, children may still have been in the process of acquiring the long passive structure in their languages. Although they should have been old enough to be able to comprehend long passives (e.g., Armon-Lotem et al., 2016), they may still have been in different stages when it comes to consolidating this knowledge, given that long passives are acquired relatively late (e.g., Bartke, 2004; Bever, 1970; Fox & Grodzinsky, 1998; Hirsch & Wexler, 2006; Verrips, 1996). This may have affected how established connections were between representational passive nodes and lexical items within and between languages (e.g., Hartsuiker & Bernolet, 2017). Given that most German-Dutch children received more input in Dutch than German, their level of acquisition and consolidation of the long passive structure in German in particular may have varied from child to child. Hence, the significant effect of relative proficiency in children for the PP-V structure might reflect the consolidation of and the connections with the PP-V structure in German (and Dutch) and, consequently, the amount of co-activation from German. In contrast, the bilingual adults were old enough to have fully acquired the long passive structure in both of their languages. Therefore, we expect that connections with the long passive structure were firmly established in all participants. If this is correct, it is not surprising that relative proficiency no longer plays a role in predicting online cross-linguistic influence in bilingual adults. In other words, because bilingual adults should have equally consolidated representations for the long passive structure in Dutch and German, we expect differences in their general proficiency levels in their languages to no longer predict the amount of structural co-activation for long passives.

With respect to our comparison of adults and children, two caveats remain. We have to keep in mind that (i) the adult participant groups were relatively small; and (ii) that the adult groups were not entirely matched on their educational level. The latter could have influenced participants' language processing abilities. Nevertheless, although the monolingually-raised participants on average had a higher educational level than the other two groups, they were not significantly faster in processing Dutch sentences. Hence, differences in education between the groups do not seem to have played an important role in our experiment.

4.6 Conclusions

In sum, in our self-paced listening task we obtained similar effects of syntactic co-activation of German during Dutch sentence processing in simultaneous bilingual adults and adolescents as we had previously observed in simultaneous bilingual children (van Dijk et al., Chapter 3), namely, (i) cross-linguistic influence was observed as a slowdown effect in online comprehension; (ii) lexical overlap, surface overlap and language dominance moderated the effect of cross-linguistic influence; and (iii) there was no evidence of processing delays due to a general effect of bilingualism. Furthermore, we found language mode, not tested in bilingual children, to further affect the effect of cross-linguistic influence in bilingual adults and adolescents. At the same time, our online results of cross-linguistic influence were less pronounced in adults and adolescents than in bilingual children. Our findings are in line with those from studies on cross-linguistic influence in simultaneous bilingual children and simultaneous and sequential bilinguals (e.g., Hopp, 2017; Hulk & Müller, 2000; Kupisch, 2012; Martohardjono et al., 2017; van Dijk et al., Chapter 2). Crucially, our findings suggest that the same mechanisms responsible for cross-linguistic influence, namely language co-activation and inhibition, play a role in bilingual adults and children.

CHAPTER 5

Cross-linguistic influence in simultaneous bilingual children's online processing of Dutch pronouns

An eye-tracking study

Abstract

In this study we investigated whether pronoun interpretation preferences from a null subject language, Turkish, influence online and offline pronoun interpretation preferences in a non-null subject language, Dutch, in Turkish-Dutch bilingual children. Furthermore, we investigated whether language dominance moderates the strength of such cross-linguistic influence. Finally, we tested whether a general bilingualism effect affects children's online and offline pronoun interpretations. We measured children's behaviour using an eye-tracking task (visual world paradigm) in combination with a picture selection task. German-Dutch bilingual and Dutch monolingual children served as control groups. We found evidence for cross-linguistic influence from Turkish in the Turkish-Dutch children's fixations when we took children's language dominance profile into account. The more balanced children were in their languages, as opposed to being Dutch-dominant, the less they fixated on the Turkish-preferred non-topic referent. We observed a similar although non-significant pattern offline. Finally, we found no evidence for a general bilingualism effect. We discuss our findings in terms of structural co-activation and inhibition.

Based on: van Dijk, C.N., Aumeistere, A., Brouwer, S., Dijkstra, T., & Unsworth, S. (unpublished manuscript). Cross-linguistic influence in simultaneous bilingual children's online processing of Dutch pronouns: an eye-tracking study.

5.1 Introduction

Bilingual children have to acquire different rules in their different languages. Consider the example in (1).

(1) Anna_i en Sophie_k leren in de bibliotheek.
Anna_i and Sophie_k study in the library
Terwijl Anna_i een boek leest, neemt $\text{zij}_{i/?k}$ een slokje water.
while Anna_i a book reads takes $\text{she}_{i/?k}$ a sip water

"Anna and Sophie are studying in the library. While Anna is reading a book, she takes a sip of water."

In order to interpret the final clause, a link has to be established between the pronoun and a referent in the discourse. In Dutch, a non-null subject language, the most likely referent would be *Anna*. In contrast, in null subject languages, such as Turkish, Italian or Greek, the referent should either be *Sophie* or an unmentioned third referent. Hence, non-null subject and null subject languages have different preferences when it comes to pronoun resolution (e.g., Ariel, 2014; Cardinaletti & Starke, 1999). It is unknown how these differences affect online pronoun resolution in bilingual children acquiring a null subject language alongside a non-null subject language. In fact, cross-linguistic influence during real-time sentence processing in general is an underexplored area in bilingual children (cf. Lemmerth & Hopp, 2019; van Dijk, Dijkstra, & Unsworth, Chapter 3). Instead, virtually all studies on cross-linguistic influence have employed elicited production tasks (e.g., Foroodi-Nejad & Paradis, 2009; Nicoladis, 2012) or offline comprehension tasks, such as forced choice (e.g., Argyri & Sorace, 2007; Serratrice, Sorace, Filiaci, & Baldo, 2009) and picture selection (e.g., Kidd, Chan, & Chiu, 2015; Serratrice, 2007).

In this study, our main aim was to deepen our knowledge of cross-linguistic influence in bilingual children during real-time sentence processing, and, in particular, during pronoun resolution by using an eye-tracking in the visual world paradigm. A secondary aim was to relate findings from this online experimental technique to children's offline interpretations on a picture-selection task. We were interested in the effect of a null subject language on pronoun resolution in a non-null subject language. We therefore tested the influence of Turkish on online and offline preferences in Dutch in Turkish-Dutch bilingual children. To our knowledge, this is the first study to use an online technique to study pronoun resolution in bilingual children.

In the remainder of this section, we discuss pronoun resolution in Dutch and Turkish in more detail. Given that language-specific literature does not always exist, especially for Turkish, our discussion focuses on the broader distinction between non-null subject languages (e.g., Dutch, English, German) and null subject languages (e.g., Turkish, Italian, Greek). We then discuss studies on cross-linguistic influence in pronoun use and offline interpretation in bilingual children and in adult second language (L2) learners (online), before formulating the hypotheses to be tested in the present study.

Pronoun resolution in non-null subject languages

In Dutch and other non-null subject languages, such as English and German, a pronoun usually refers back to the most accessible referent in the discourse (e.g., Ariel, 1994, 2014; Cardinaletti & Starke, 1999). Accessibility depends on various factors, such as the recency, grammatical role, and topicality of a referent (e.g., Ariel, 2014; Givón, 1983; Järvikivi, Pyykkönen-Klauck, Schimke, Colonna, & Hemforth, 2014; Song & Fisher, 2005). Referents that have been mentioned recently, that are the subject of the (preceding) sentence and that are the topic of the discourse are typically more prominent in the discourse and therefore more accessible as antecedents. In (1), this would make *Anna* the most likely antecedent of the pronoun *zij* ("she").

Sources of information other than discourse status guide pronoun resolution as well. These include, for example, syntactic information such as gender and number information on the pronoun (e.g., Arnold et al., 2000; Tyler, 1983). Importantly, syntactic information can override the discourse preference to bind a pronoun to the most accessible referent in the discourse. For example, in (2) the pronoun is disambiguated by gender information. The masculine pronoun *hij* ("he") can only refer to the disjoint referent *Thomas* and not to the local referent *Anna*.

(2) Anna$_i$ en Thomas$_k$ leren in de bibliotheek.
Anna$_i$ and Thomas$_k$ study in the library
Terwijl Anna$_i$ een boek leest, neemt hij$_{*i/k}$ een slokje water.
while Anna$_i$ a book reads takes he$_{*i/k}$ a sip water

"Anna and Thomas are studying in the library. While Anna is reading a book, he takes a sip of water."

Monolingual adult speakers of non-null subject languages rapidly integrate information during pronoun resolution about the accessibility status of the referents and gender (e.g., Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Hemforth et al., 2010). Online experiments have also shown that young monolingual children are already sensitive to the discourse status of referents by the age of 3;0 (e.g., Järvikivi et al., 2014; Song & Fisher, 2005; Tyler, 1983). At the same time, however, evidence suggests that the integration of discourse information in children is relatively slow (e.g., Arnold, Brown-Schmidt, & Trueswell, 2007; Järvikivi et al., 2014). In particular, whilst adults have been found to integrate information about the discourse status of the referents directly or shortly after hearing the pronoun (e.g., Arnold et al., 2000; Hemforth et al., 2010; Järvikivi, van Gompel, Hyönä, & Bertram, 2005), children do not use discourse information online until much later during sentence processing (e.g., Arnold et al., 2007; Song & Fisher, 2005). Conversely, monolingual children have been found to rapidly integrate gender information, showing quantitatively similar patterns to monolingual adults by the age of 5;0 (e.g., Arnold et al., 2007).

Pronoun resolution in null subject languages

Null subject languages, such as Turkish, Italian and Greek, allow both overt pronouns and null pronouns in subject positions (e.g., Azar & Özyürek, 2015; Cardinaletti & Starke, 1999; Carminati, 2002). A null pronoun typically refers back to the most accessible referent in the discourse (e.g., Cardinaletti & Starke, 1999). Hence, to refer back to Anna in the Turkish example in (3), a null pronoun would be preferred.

(3) Anna$_i$ kitap okurken, Ø$_i$ sudan bir yudum alıyor.
Anna$_i$ book read Ø$_i$ water a sip take

"While Anna is reading a book, she takes a sip of water."

The use of overt pronouns in Turkish is marked. Overt pronouns usually signal a shift in topic or place emphasis on their antecedent (e.g., Azar & Özyürek, 2015; Azar, Özyürek, & Backus, 2020; Enç, 1986). Consequently, in the Turkish translation of (1) given in (4), it would be pragmatically infelicitous for the Turkish overt pronoun *o* to refer back to the topic of the discourse, which is *Anna* (4). Instead, it is linked to a non-topic antecedent, which is either *Sophie* or an unmentioned third person. Whilst Turkish pronouns are marked for person and number, they do not carry grammatical gender information. Because of this, gender cannot be used as a disambiguating cue in Turkish pronoun resolution.

(4) $Anna_i$ ve $Sophie_k$ kütüphane çalışıyorlar.
$Anna_i$ and $Sophie_k$ library work
$Anna_i$ kitap okurken, $o_{?i/k}$ sudan bir yudum alıyor.
$Anna_i$ book read s/$he_{?i/k}$ water a sip take

"Anna and Sophie are studying in the library. While Anna is reading a book, she takes a sip of water."

Online studies with adult native speakers of null subject languages show that adults have a topic antecedent bias during real-time processing of null pronouns but not during real-time processing of overt pronouns (e.g., Carminati, 2002; Filiaci, 2010; Papadopoulou, Peristeri, Plemenou, Marinis, & Tsimpli, 2015). Studies have also shown that monolingual children show a similar bias in their offline interpretations of null and overt pronouns as adults (e.g., Papadopoulou et al., 2015; Serratrice, 2007). To be more precise, children prefer to bind a null subject pronoun to the topic of the discourse and an overt pronoun to a non-discourse topic. Offline, however, this preference is less strong than in adults, even at the age of 8;0 (e.g., Argyri & Sorace, 2007; Papadopoulou et al., 2015; Serratrice, 2007). In contrast, 6-year-old monolingual Greek children display a non-topic antecedent bias *during* real-time processing of overt pronouns quantitatively similar to adults. These findings suggest that children acquiring a null subject language can employ discourse information during pronoun resolution (Papadopoulou et al., 2015). However, they (sometimes) fail to integrate this information in their offline interpretations.

To sum up, speakers of non-null subject and null subject languages have different preferences when it comes to the resolution of overt pronouns. It takes time before monolingual children are able to use these settings in their online pronoun resolution in an adultlike manner. To our knowledge, there are no studies that have investigated online pronoun resolution in bilingual children. However, a number of studies have examined how the parallel acquisition of a non-null subject and a null subject language affects children's pronoun use and their *offline* pronoun interpretations. We now turn to these studies.

Cross-linguistic influence in bilingual children's pronoun choice and interpretation

Cross-linguistic influence in bilingual children's pronoun choices has been found to occur unidirectionally from the non-null subject language into the null subject language. Evidence comes from spontaneous speech production in younger bilingual children and from forced choice tasks in older bilingual children (e.g., Argyri & Sorace, 2007; Schmitz, Patuto, & Müller, 2011; Serratrice, Sorace, & Paoli, 2004; Sorace, Serratrice, Filiaci, & Baldo, 2009). In spontaneous speech studies, it has been observed that children between 2 and 5 years of age overproduce overt subjects in null subject languages such as Hebrew, Italian and Turkish compared to monolingual peers (e.g., Hacohen & Schaeffer, 2007; Haznedar, 2007; Schmitz et al., 2011; Serratrice et al., 2004). These children were acquiring a non-null subject language – typically English – alongside their null subject language. In contrast, there is no evidence in favour of cross-linguistic influence in the opposite direction: bilingual children did not omit the subject pronoun more often or in different contexts in their non-null subject language than monolingual peers (e.g., Schmitz et al., 2011; Serratrice et al., 2004).

Pronoun choices have also been investigated in older bilingual children. Argyri and Sorace (2007) tested 7-to-9-year-old Greek-English bilingual children. They found no evidence for cross-linguistic influence in children's elicited speech production in either language. However, in a forced choice task, unidirectional cross-linguistic influence from English into Greek was observed. Bilingual children chose sentences with a pragmatically infelicitous overt subject pronoun more often than monolingual peers. No effects were observed in the opposite direction. In other words, bilingual children did not choose sentences with a null subject significantly more often in English than monolingual peers. Similar effects were obtained in a study with Italian-English bilingual children aged between 6 and 7 years (Sorace et al., 2009). Effects in both studies were modulated by children's language dominance profiles as measured by the language spoken in children's environment (English in the UK versus Italian in Italy). Evidence for cross-linguistic influence was only attested in children living in the UK.

Evidence for cross-linguistic influence in bilingual children's pronoun comprehension was reported in Serratrice's (2007) study of Italian pronoun interpretation in Italian-English bilingual children. Using data from a picture selection task, she found that bilingual children were more likely than monolingual children and adults to choose the subject antecedent of an overt pronoun rather than the object antecedent.

Cross-linguistic influence in children's pronoun choices and interpretations has been explained in terms of co-activation and priming during language processing (e.g., Serratrice, 2007, 2016; Sorace & Serratrice, 2009; Sorace et al., 2009, also see Nicoladis, 2006, 2012; Nicoladis, Rose, & Foursha-Stevenson, 2010, for a similar claim for word order phenomena). According to such accounts, pronoun preferences in bilingual children's one language compete for activation when children use or interpret a pronoun in their other language. This can sometimes result in children selecting a pronoun structure from the language not in use (e.g., Sorace & Serratrice, 2009). It has furthermore been argued that the use of overt pronouns in one language can prime the use of overt pronouns in similar contexts over time. Consequently, English-Italian bilingual children who, for instance, receive relatively more input in English, might develop English pronoun preferences in their less frequent language, Italian.

There is an alternative explanation for differences in pronoun choices between bilingual children and their monolingual peers (e.g., Sorace, 2011; Sorace & Serratrice, 2009; Sorace et al., 2009). Bilingual children might be less accurate in their pronoun use irrespective of the pronoun properties of their other language, but due to a *general bilingualism* effect. Sorace and colleagues (2009) tested pronoun choices in Italian in a group of Italian-Spanish bilingual children in addition to the Italian-English bilingual children in their study. Spanish, like Italian, is a null subject language. The authors, therefore, expected the Italian-Spanish children to prefer sentences with a null subject pronoun to refer to the topic of the discourse. Nevertheless, the Italian-Spanish bilingual children chose the pragmatically infelicitous overt pronoun more often to refer to the topic of the discourse than their Italian monolingual peers. This finding could not be explained in terms of cross-linguistic influence. Instead, Sorace and others (Sorace, 2011; Sorace & Serratrice, 2009; Sorace et al., 2009) argued that differences in pronoun choices between bilingual and monolingual children could be – at least partially – explained by general processing difficulties in bilingual children. In particular, they suggested that bilingual children might have insufficient processing resources available to integrate discourse information during pronoun resolution. As a consequence, they fall back on a default strategy – i.e., the use of an overt pronoun to establish reference in the discourse.

It is unclear from the existing studies with bilingual children how cross-linguistic influence and/or general processing difficulties may affect online pronoun resolution. Furthermore, to the best of our knowledge, the effect of a null subject language on pronoun *comprehension* in a non-null subject language – the direction investigated in this study – has not yet been explored. In contrast, online techniques have been employed to test for cross-linguistic influence during pronoun resolution in this direction in studies with adult L2 learners. We now discuss findings from these studies.

Cross-linguistic influence in pronoun resolution in adult L2 learners
Our study was inspired by the few available studies that have compared offline and online pronoun resolution in adult L2 learners of a non-null subject language with a L1 null subject language (e.g., Cunnings, Fotiadou, & Tsimpli, 2017; Roberts, Gullberg, & Indefrey, 2008; Schimke, de la Fuente, Hemforth, & Colonna, 2018). These studies show that L2 pronoun resolution can be affected by cross-linguistic influence and general processing difficulties, although results and the interpretation thereof differ from study to study. We discuss the three studies we based our design on in more detail.

The first study was concerned with our language combination of interest, Turkish and Dutch. More specifically, Roberts and colleagues (2008) investigated the influence of Turkish as a first language (L1) on offline and online pronoun resolution in Dutch as L2. Adult native speakers of Dutch and L2 learners of Dutch with German as L1 served as control groups. Because Dutch and German are both non-null subject languages, similar behaviour was expected in the German-Dutch and the L1 Dutch group. The authors tested participants' interpretations of pronouns occurring in three different contexts: local, disjoint and optional. In the local and the disjoint contexts, pronouns were disambiguated by number cues and referred either to a local referent or to a disjoint referent. The local referent was mentioned in the clause directly preceding the pronoun and was therefore most prominent in the discourse. The disjoint referent was mentioned earlier in the discourse and therefore less prominent. In the optional context, the pronoun was ambiguous and was grammatically congruent with both a local or disjoint interpretation.

In an offline task prompting the meaning of the pronoun, all groups (almost) always chose the local antecedent in the local condition and the disjoint antecedent in the disjoint condition. Furthermore, the Dutch L1 and the German-Dutch participants chose the local referent in the optional condition more than 90% of the time. This clear local bias reflects the preference to link a pronoun to the most accessible referent in the discourse in Dutch and German. In contrast, the Turkish-Dutch participants chose the disjoint referent as pronoun antecedent only about 50% of the time in the optional condition. This suggests that Turkish as an L1 influences offline pronoun interpretation in Dutch as an L2.

In an online task, the authors further observed that the Dutch L1 group was significantly faster to read the verb preceding the pronoun and the pronoun itself in the optional condition than in the local and disjoint conditions. This suggested that pronoun resolution in native speakers was facilitated when they only had to integrate the discourse status of the possible referents and not their number features. In contrast, the Turkish-Dutch group slowed down in the optional condition relative to the other two conditions. This could not be interpreted as cross-linguistic influence from Turkish, however, because the German-Dutch group showed a similar effect. Instead, Roberts and colleagues accounted for the online behaviour of the two bilingual groups in terms of general processing difficulties. Their argument was as follows: in the optional condition the meaning of the pronoun was ambiguous. Participants could only resolve this ambiguity by taking into account the discourse status of the two referents. However, according to the authors, L2 learners had difficulties integrating discourse information during real-time processing due to a general effect of bilingualism (also see e.g., Clahsen & Felser, 2006; Sorace, 2011). This explained their slower reading times.

Cunnings, Fotiadou and Tsimpli (2017) explored Roberts et al.'s (2008) findings in more detail. They used the visual world paradigm combined with a picture selection task to test pronoun resolution in English. Participants were adult native speakers of English and L2 learners of English with Greek as L1. In half of the sentences, pronouns were disambiguated by gender, either matching the subject or the object of the previous sentence. In the other half of the sentences, the pronoun could initially refer back to both referents. Later during the sentence, the pronoun was disambiguated by lexical information. L2 learners behaved similarly to English native speakers in the gender-disambiguated pronoun condition. This was reflected by accurate offline referent choices and rapid online integration of gender information on the pronoun. In the ambiguous conditions the L2 and native group showed an initial preference for the subject antecedent after having heard the pronoun similar to the native group – as evidenced by more looks at the subject of the previous clause. Hence, in contrast to Roberts et al. (2008), the authors did not observe L2 difficulties with integrating discourse information online. However, the L2 learners had more difficulties reanalysing their initial preference, when the following noun disambiguated the pronoun towards the object referent. This was reflected by more online looks to and more offline choices of the subject referent in the L2 group compared to the English L1 group. Cross-linguistic influence from Greek could not account for the offline differences between the English L1 speakers and the L2 speakers. Instead, Cunnings and colleagues suggested that L2 learners have difficulties reanalysing their initial pronoun interpretation, irrespective of their L1.

Finally, Schimke and colleagues (2018) compared offline and online pronoun resolution in German as L2 in a group of L1 Spanish and L1 French speakers. Spanish is a null subject language and French is a non-null subject language. To reduce processing demands they also used the visual world paradigm. The pronouns in their stimuli could either refer back to the subject or the object of the previous clause. By the end of the sentence, disambiguating information was given about the pronoun. The French-German participants chose the subject of the previous clause more often as pronoun antecedent (80.4%) than the Spanish-German participants (66.7%). This finding was consistent with cross-linguistic influence. The difference between the two groups was not significant, however. Online, the Spanish-German participants did behave significantly different from the French-German group. First of all, after having encountered the pronoun, the French-German participants looked more at the subject of the previous clause than the Spanish-German participants. Second, the number of looks to the subject in the French-German group differed significantly from chance, whereas it did not in the Spanish-German group. The authors concluded that L2 learners' L1 influenced their online pronoun resolution in their L2. Furthermore, there was a possible L1 effect in participants' offline responses as well. The authors argued that L1 effects were less pronounced offline due to an interaction with a more general bilingualism effect. In particular, the authors suggested that due to the conscious decision involved in an offline task, L2 learners might have opted for the subject interpretation as a default strategy (also see Kaiser, 2011).

In sum, studies with adult L2 learners show that a null subject language can influence online pronoun resolution in a non-null subject language. However, there is also evidence for general processing difficulties in L2 learners. Furthermore, offline patterns are not necessarily reflected online.

5.2 The present study

In the present study, we addressed whether effects of cross-linguistic influence and general processing difficulties were present in online and offline pronoun resolution in Turkish-Dutch children. In order to do so, we compared their online and offline behaviour to a group of German-Dutch and a group of Dutch monolingual peers, following Roberts et al. (2008). Because Dutch and German have similar preferences regarding pronoun resolution, we expected no effects of cross-linguistic influence in the German-Dutch group (e.g., Roberts et al., 2008). Furthermore, a general effect of bilingualism was ruled out in the monolingual group.

In line with Cunnings et al. (2017) and Schimke et al. (2018), we used an eye-tracking task (visual world paradigm) that we combined with an offline picture selection task (e.g., Cunnings et al., 2017). We investigated children's fixations while they were listening to *while*-clauses (adapted from Roberts et al., 2008) and looking at pictures of two possible referents on a screen (see Figure 5.1). Sentences were either disambiguated by gender information (5 and 6) or ambiguous between a local and a disjoint reading (7).

(5) **Local**

Anna$_i$ en Thomas$_k$ leren in de bibliotheek.
Anna$_i$ and Thomas$_k$ study in the library
Terwijl Anna$_i$ een boek leest, neemt zij$_{i/*k}$ een slokje water.
while Anna$_i$ a book reads takes she$_{i/*k}$ a sip water

"Anna and Thomas are studying in the library. While Anna is reading a book, she takes a sip of water."

(6) **Disjoint**

Anna$_i$ en Thomas$_k$ leren in de bibliotheek.
Anna$_i$ and Thomas$_k$ study in the library
Terwijl Anna$_i$ een boek leest, neemt hij$_{*i/k}$ een slokje water.
while Anna$_i$ a book reads takes he$_{*i/k}$ a sip water

"Anna and Thomas are studying in the library. While Anna is reading a book, he takes a sip of water."

(7) **Optional**

Anna$_i$ en Sophie$_k$ leren in de bibliotheek.
Anna$_i$ and Sophie$_k$ study in the library
Terwijl Anna$_i$ een boek leest, neemt zij$_{i/k}$ een slokje water.
while Anna$_i$ a book reads takes she$_{i/k}$ a sip water

"Anna and Sophie are studying in the library. While Anna is reading a book, she takes a sip of water."

We tested four hypotheses. First, we hypothesized that pronoun resolution preferences of bilingual children's one language can influence online pronoun resolution in another language, in line with offline findings in bilingual children (e.g., Serratrice, 2007) and online findings in adult L2 learners (e.g., Schimke et al., 2018). In particular, we expected Turkish-Dutch children to look more at the disjoint referent and less at the local referent when listening to Dutch pronouns compared to Dutch monolingual and German-Dutch children.

Second, we hypothesized that effects of online cross-linguistic influence would become stronger in bilingual children with increased dominance in Turkish, in line with previous studies on bilingual children's pronoun choices (e.g., Argyri & Sorace, 2007; Sorace et al., 2009). Specifically, we predicted more looks to the disjoint referent and less looks to the local referent the more Turkish-dominant children were.

Third, we hypothesized that online pronoun resolution in both Turkish-Dutch and German-Dutch bilingual children might be affected by general processing difficulties. This has been observed in offline pronoun interpretations of simultaneous bilingual children (e.g., Sorace et al., 2009) and offline and online pronoun interpretations in adult L2 learners (e.g., Cunnings et al., 2017; Roberts et al., 2008; Schimke et al., 2018). We expected such general processing difficulties to be reflected by more fixations on the default local referent in the bilingual groups than in the monolingual group, in line with findings for L2 learners (e.g., Cunnings et al., 2017).

Fourth, we hypothesized that if cross-linguistic influence during pronoun resolution is the result of language co-activation and priming during sentence processing (e.g., Serratrice, 2007; Serratrice, 2016; Sorace & Serratrice, 2009), offline cross-linguistic influence might be less pronounced than online cross-linguistic influence. Only when effects of cross-linguistic influence during real-time pronoun resolution are strong enough, we expect them to also be visible in children's offline referent choices. Hence, whilst we did not rule out that Turkish-Dutch children would choose the disjoint referent as the pronoun antecedent more often than the German-Dutch bilingual and Dutch monolingual children in the offline task, we expected this difference to be more pronounced in the online task. Similarly, any relation between cross-linguistic influence and language dominance was expected to be less pronounced offline than online.

5.3 Method

Participants

A total of 54 children were tested: 17 Turkish-Dutch bilingual children (mean age = 9.1; *SD* = 1.2), 22 German-Dutch bilingual children (mean age = 8.8; *SD*

= 1.2) and 14 Dutch monolingual children (mean age = 8.2; *SD* = 0.9).[1] All children were between 7 and 10 years old and were living in the Netherlands during time of testing. Bilingual children were first exposed to Turkish or German at birth. Their age of onset to Dutch was at birth as well or maximally 8 months thereafter for 14 Turkish-Dutch children and 21 German-Dutch children. The remaining 4 children had started to acquire Dutch after the age of 1;0 but before the age of 3;0.

The bilingual children's parents were interviewed using an extensive questionnaire (*Bilingual Language Exposure Calculator*; Unsworth, 2013) to assess children's language experience in both of their languages. A summary is provided in Table 5.1. Children's cumulative input shows the percentage of input children had received since birth in Turkish and German relative to Dutch. On average, both groups had received less input in Turkish/German than Dutch. However, there was quite some individual variation within the groups. Similarly, children were on average more exposed to Dutch at the time of testing than to Turkish and German (*Current input*). Again, there was considerable variation, although standard deviations and ranges show that most children received more input in Dutch than in their other language. Children's cumulative and current input served as our first two measures of language dominance (following Unsworth, Chondrogianni, & Skarabela, 2018; van Dijk, Dijkstra, et al., Chapter 3).

Table 5.1. Overview of background variables for bilingual children (means, standard deviations and ranges) and independent t-tests.

		Turkish-Dutch	**German-Dutch**	**t-test**
Cumulative input	Turkish/ German (%)	39.2 (13.7) 12.9–67.7	43.5 (11.1) 22.2–65.8	$t(28.4) = 1.0$; $p = .321$
Current input	Turkish/ German (%)	28.9 (8.8) 12.4–42.3	34.6 (14.1) 9.4–59.2	$t(33.9) = 1.5$; $p = .143$
Parental educational level	Average mother & father	4.4 (1.3) 2.5–6.5	6.3 (1.4) 3–7.5	$t(33.2) = 4.3$; $p < .001$

[1] Group sizes are smaller than intended because testing had to be postponed due to Covid-19.

Tasks

Eye-tracking task

Materials. We created 36 triplets of short stories such as those in (5) through (7). Characters were chosen from a set of 6 characters: three girls (*Anna, Sophie* and *Lieke*) and three boys (*Thomas, Joris* and *Peter*). Half of the stories contained the feminine pronoun *zij* ("she") and half the masculine pronoun *hij* ("he"). We chose to use strong pronouns as opposed to the weak pronouns *ze* ("she") and *ie* ("he") to increase the pragmatic acceptability of a disjoint reading (e.g., Kaiser, 2011). Each story was followed by a question targeting the pronoun (e.g., Wie nam een slokje water? "Who took a sip of water?"). This served to measure children's offline interpretation of the pronoun. For a complete list of stories and questions, see the appendix (A5.1).

The combination and order of the two characters in each story and the subject of the experimental sentence were counter-balanced between stories. All sentences were recorded by a female native speaker of Dutch using neutral intonation. After recording, we aligned the onset of the subordinate clause, the main clause containing the pronoun, the onset of the pronoun and the onset of material after the pronoun in the different stories in Praat (Boersma & Weenink, 2018). The onset of the subordinate clause was set to 4108 ms (*SD* = 73.4), the onset of the main clause to 7066 ms (*SD* = 3.6), the onset of the pronoun to 7677 ms (*SD* = 12.6), and following material to 8217 ms (*SD* = 16.6).

Three different lists were created with every story appearing only in one version. Every list contained 12 optional, 12 local and 12 disjoint sentences. The order of the stories in each list was then pseudo-randomized such that each child in the same language group listened to a different list.

The experimental items were interspersed with 36 fillers items with different characters. These items consisted of 9 possessive structures, 9 dative structures, 9 long passives, and 9 encouragements. Except for the encouragements, filler items were followed by a comprehension question to make sure that children were paying attention to the task.

Every story was accompanied by the pictures of the two characters involved: one at the left side and one at the right side of a laptop screen. Their order was counter-balanced between items. The names of the six child characters were written on their clothes. The characters in the filler items were identifiable by their clothes. A picture of a third object mentioned in the stories was displayed underneath the characters. This was the object mentioned in the subordinate clause in the experimental items (see Figure 5.1).

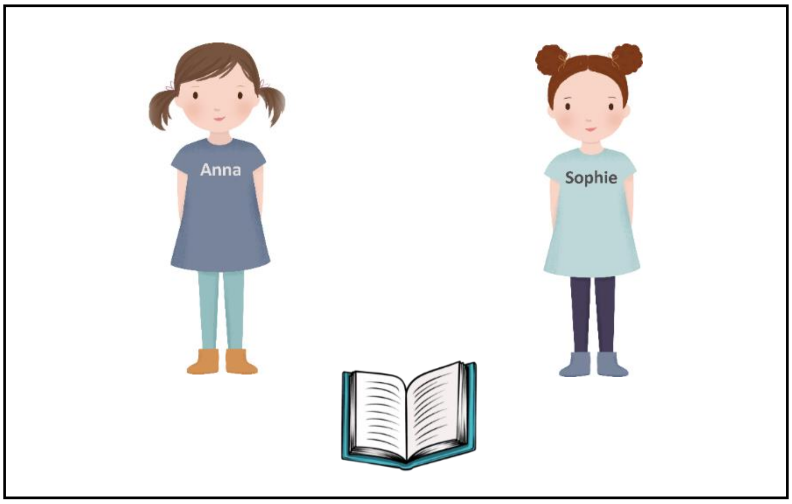


Figure 5.1. Example of visual stimuli in eye-tracking task in optional condition (see 7, repeated in 8).

(8) Optional

Anna$_i$ en Sophie$_k$ leren in de bibliotheek. Terwijl Anna$_i$ een boek leest, neemt zij$_{i/k}$
Anna$_i$ and Sophie$_k$ study in the library while Anna$_i$ a book reads takes she$_{i/k}$
een slokje water.
a sip water
"Anna and Sophie are studying in the library. While Anna is reading a book, she takes a sip of water."

Procedure. Children were sitting in front of a laptop and were wearing headphones. At the start of the task, they were introduced to the six children in the experiment. To make sure that they were able to read the characters' names on their shirts, they were asked to name each out loud. They were then told they were going to hear stories about these characters and that they would be asked a question after each story. In order to answer the questions, they had to press one of two large buttons in front of them (corresponding to the location of the characters on the screen). The tester also explained to them that their eye-movements would be measured and that it was therefore important that they sit as still as possible and keep their eyes on the screen during the entire task.

First, children received four practice items. One practice item contained a pronoun that had two possible referents. Another practice item depicted a situation in which the same character performed two actions at the same time. This accustomed children with the different possible interpretations of the experimental items. When necessary, the tester gave feedback.

During the experimental phase children listened to 72 stories (36 experimental, 36 fillers). Children's eye-movements were recorded during listening by a Tobii pro camera (120 Hz) that was attached below the laptop screen. At the start of the experiment a calibration procedure was used to control for drifts. At the start of each trial, children saw three pictures on the screen. The spoken stimuli were started at the same time the pictures were made visible. Experimental stories lasted between 9.223 ms and 10.781 ms. After each story the pictures remained on the screen for 750 ms. Then children heard a beep followed by a question about the story. During the question, the third object was removed from the screen. Children had to answer the questions by selecting one of the two characters on the screen. After every eight items, children received an encouragement and were given the possibility to take a break. The task took about 40 minutes to complete.

Scoring & data preparation. To code children's eye-tracking data for the experimental items, we divided the laptop screen into three areas of interest: local character, disjoint character, and distractor (third object). For each 8 ms time frame we coded which area of interest children had been fixating at. We then aggregated the 8 ms time windows in 40 ms time bins for each child for each trial. Per time bin we calculated the proportion of looks to the three areas of interest. Proportions were transformed using the empirical logit transformation (e.g., Barr, 2008). For each time bin we then calculated the difference between children's fixations on the local referent and the disjoint referent by subtracting the empirical logit-transformed fixation proportions at the disjoint referent from those at the local referent (local-disjoint). This difference between proportions was our dependent variable in the analyses.

Furthermore, we defined two time windows of interest: (i) fixations while listening to the pronoun; and (ii) fixations until 2000 ms after the pronoun offset. Because previous studies with monolingual children have shown that integration of discourse information can occur late during pronoun resolution (e.g., Arnold et al., 2007; Song & Fisher, 2005), we used the two time windows to explore initial and later preferences for the local and disjoint referent online. Assuming that planning a saccade takes at least 200 ms (e.g., Järvikivi et al., 2014; Matin, Shao, & Boff, 1993), we set the first time window from the onset of the pronoun until 200 ms after the offset of the pronoun (0 ms – 740 ms), and the second time window from 740 ms until 2000 ms.

For children's offline referent choices, button responses were coded as 1 (local response) or 0 (disjoint response) for each item.

Turkish pronoun task

The Dutch eye-tracking task was translated into Turkish and administered as an offline task during a second test session. This allowed us to confirm whether the Turkish-Dutch children were aware of the discourse properties of Turkish pronouns. For reasons of space, the task itself and its results are discussed in A5.2 in the appendix.

Cross-linguistic lexical tasks (CLTs)

To assess children's lexical proficiency in their languages, we measured their vocabulary production skills using the CLTs in Dutch, German and Turkish from the LITMUS-battery (LITMUS-CLT: Haman, Łuniewska & Pomiechowska, 2015; Rinker & Gagarina, 2017; Ünal, Tunçer & Ege, 2012; van Wonderen et al., 2017). Every task consisted of 60 coloured pictures (30 nouns and 30 verbs), which children had to name. The tasks were constructed in such a way that the lexical items were comparable in terms of complexity and age of onset of acquisition across languages (Haman et al., 2015).

For each CLT we calculated for each child the percentage of items that were named correctly. In deciding whether or not children's responses were accurate, we followed the scoring procedure by Bohnacker, Lindgren and Oztekin (2016). In order to use Turkish-Dutch children's CLT scores as our third measure of dominance, we subtracted children's Dutch CLT scores from their Turkish CLT scores (following Yip & Matthews, 2006). A score of zero reflected equal proficiency in bilingual children's languages. A negative difference score reflected a better performance on the Dutch CLT and a positive difference score a better performance on the Turkish CLT. We interpret children's difference score as a proxy of their relative proficiency in their languages.

Digit span task

Children's short-term and working memory capacity was measured by a digit span task from the Automated Working Memory Assessment (AWMA; Alloway, 2012). The task consisted of one block during which children had to repeat sequences of digits in the same order as they had heard them (forward:

proxy of short-term memory capacity) and of one block during which children had to repeat sequences of digits in the reversed order (backward: proxy of working memory capacity). Sequences of digits ranged from 1 to 8 in the forward block (max. 48 points) and from 2 to 7 in the backward block (max. 36 points). When children incorrectly repeated three sequences of digits with the same length the block was terminated.

Procedure

The bilingual children were tested at home during two sessions by trained testers. For 3 Turkish-Dutch children and 5 German-Dutch children the second test session took place online, due to Covid-19. For the same reason, the second test session for one German-Dutch and one Turkish-Dutch child had to be cancelled completely. Unfortunately, we could not calculate dominance scores for these children, so they were excluded from the dominance analyses (but not from the group analyses). The first session was in Dutch and the second session in German or Turkish. Testers were (near-)native speakers of the language of testing. In the Dutch test session, children first conducted the eye-tracking task, then the Dutch CLT and then the digit span task. This session lasted approximately 75 minutes. In the second test session, the Turkish-Dutch children conducted the Turkish pronoun task and the Turkish CLT (45 minutes). The German-Dutch children conducted the German CLT (15 minutes). Families were rewarded with a €15,- voucher for their participation. Monolingual Dutch children participated in the monolingual test session only and were either tested in a quiet room at school or at the university. They received a small gift for their participation. Parents gave written consent for their children to participate.

5.4 Results

Background variables

Table 5.2 shows children's scores on the background measures. The Turkish-Dutch children performed relatively well on the Dutch CLT, with an average score exceeding 80%. Children's score on the Turkish CLT was lower. Furthermore, variation in children's CLT scores was much larger in the Turkish than the Dutch task. The average difference score and its range reflected (i) an overall better performance on the Dutch than the Turkish CLT; and (ii) that children's relative lexical proficiency in their languages ranged from Dutch-dominant to balanced, rather than Turkish-dominant. There was a moderate significant correlation between children's Dutch CLT performance and their difference score (Pearson's correlation: $r(16) = -.52$; $p = .037$), and a strong significant correlation between children's Turkish CLT score and their difference score (Pearson's correlation: $r(16) = 0.94$; $p < .001$).

Furthermore, the German-Dutch and Dutch monolingual children performed significantly better on the Dutch CLT than the Turkish-Dutch children, as assessed with a one-way ANOVA and post-hoc tests (Tukey correction). The German-Dutch bilingual children also had a significantly higher score on the German CLT than the Turkish-Dutch bilingual children had on the Turkish CLT, as assessed with a t-test. There were no significant differences between groups on the digit span tasks.

Table 5.2. Scores on background measures for monolingual and bilingual children (means, standard deviations, and ranges).

		Turkish-Dutch	**German-Dutch**	**Dutch**	**Test statistic**
CLT	Dutch (%)	81.7 (8.5)[a] 68–95	92.6 (6.5)[b] 67–100	89.4 (5.2)[b] 82–100	$F(2, 50) = 12.1$; $p < .001$
	Turkish/ German (%)	61.1 (20.9)[a] 14–88	82.5 (13.7)[b] 43–97		$t(24.5) = 3.6$; $p = .002$
	Difference score	19.8 (23.9) -71–12			
Digit span task	Forward	24.6 (4.6) 19–38	25.3 (3.7) 19–31	24.2 (3.1) 21–32	$F(2, 50) = 0.3$; $p = .713$
	Backward	11.8 (4.1) 6–23	11.8 (3.3) 7–18	10.4 (3.0) 7–19	$F(2, 50) = 0.8$ $p = .462$

Note One-way ANOVAs for comparisons between three participant groups and independent t-tests for comparisons between two groups. For each variable, scores differing significantly between groups are indicated with superscript. Within each row, different superscripts denote significant differences between groups.

Online pronoun resolution (eye-tracking data)

Figures 5.2, 5.3 and 5.4 shows the average proportion of children's fixations on the local and disjoint character over time in the local, disjoint and optional conditions. The time window plotted is -200 ms before and 2000 ms after the pronoun onset. The dotted vertical lines indicate the average pronoun onset and offset and the solid vertical line divides the pronoun and the post-pronoun window.

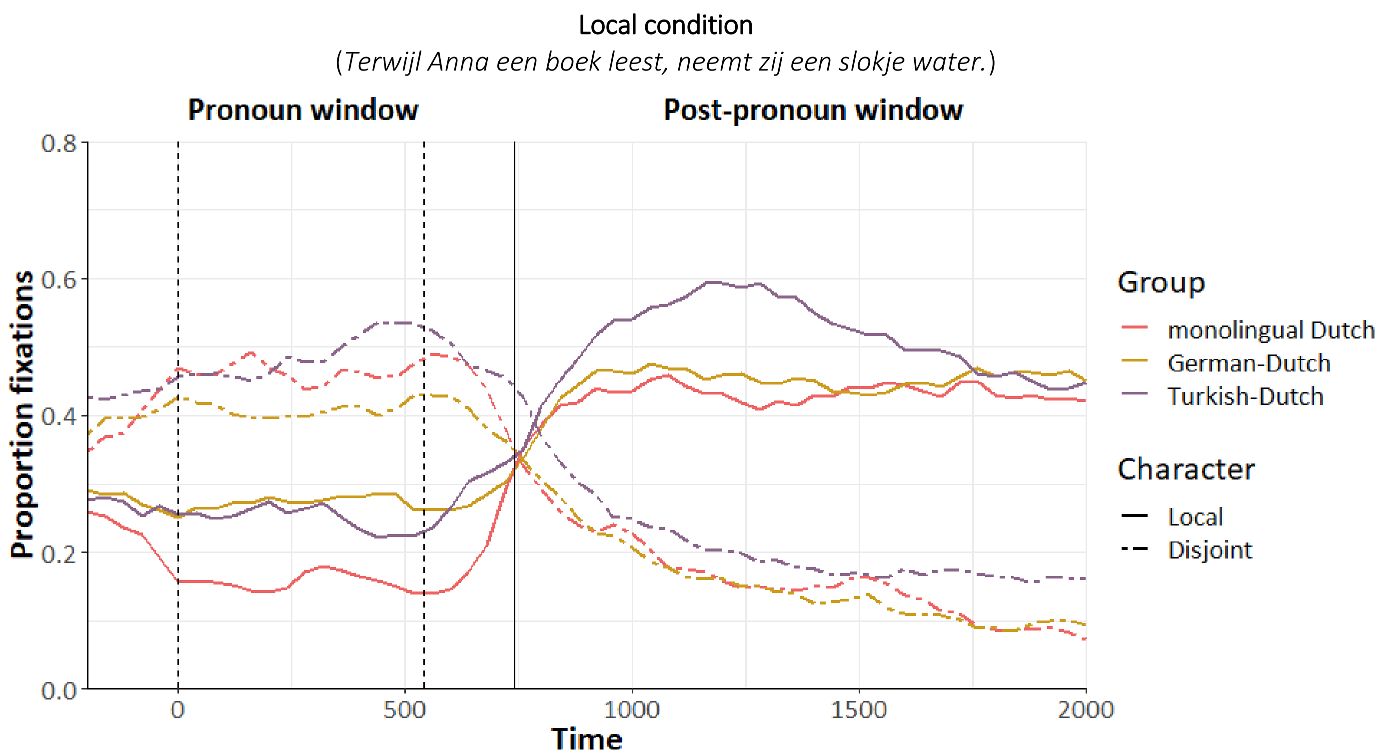


Figure 5.2. Average proportion of fixations on the local and disjoint characters in the local condition for the three participant groups. The first vertical line indicates the average pronoun onset and the second vertical line indicates the average pronoun offset.

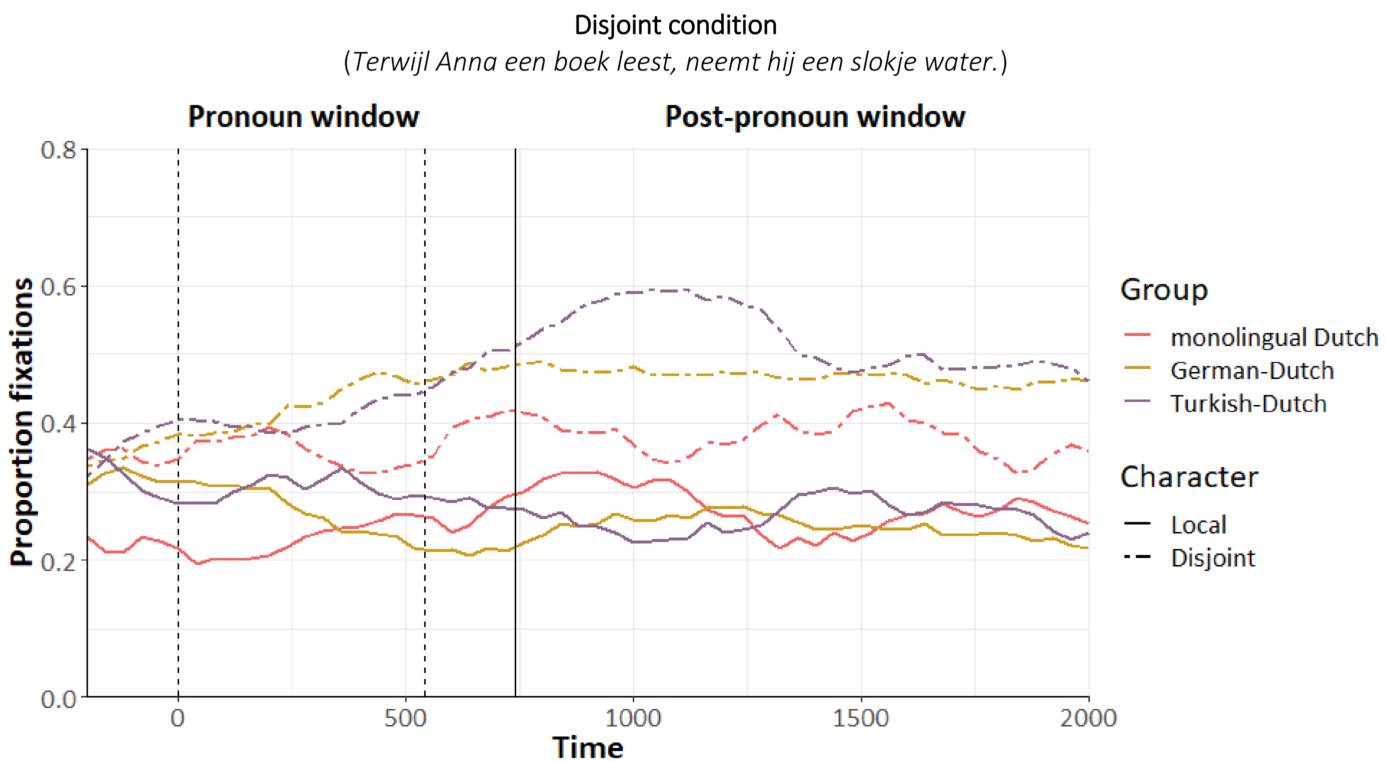


Figure 5.3. Average proportion of fixations on the local and disjoint characters in the disjoint condition for the three participant groups. The first vertical line indicates the average pronoun onset and the second vertical line indicates the average pronoun offset.

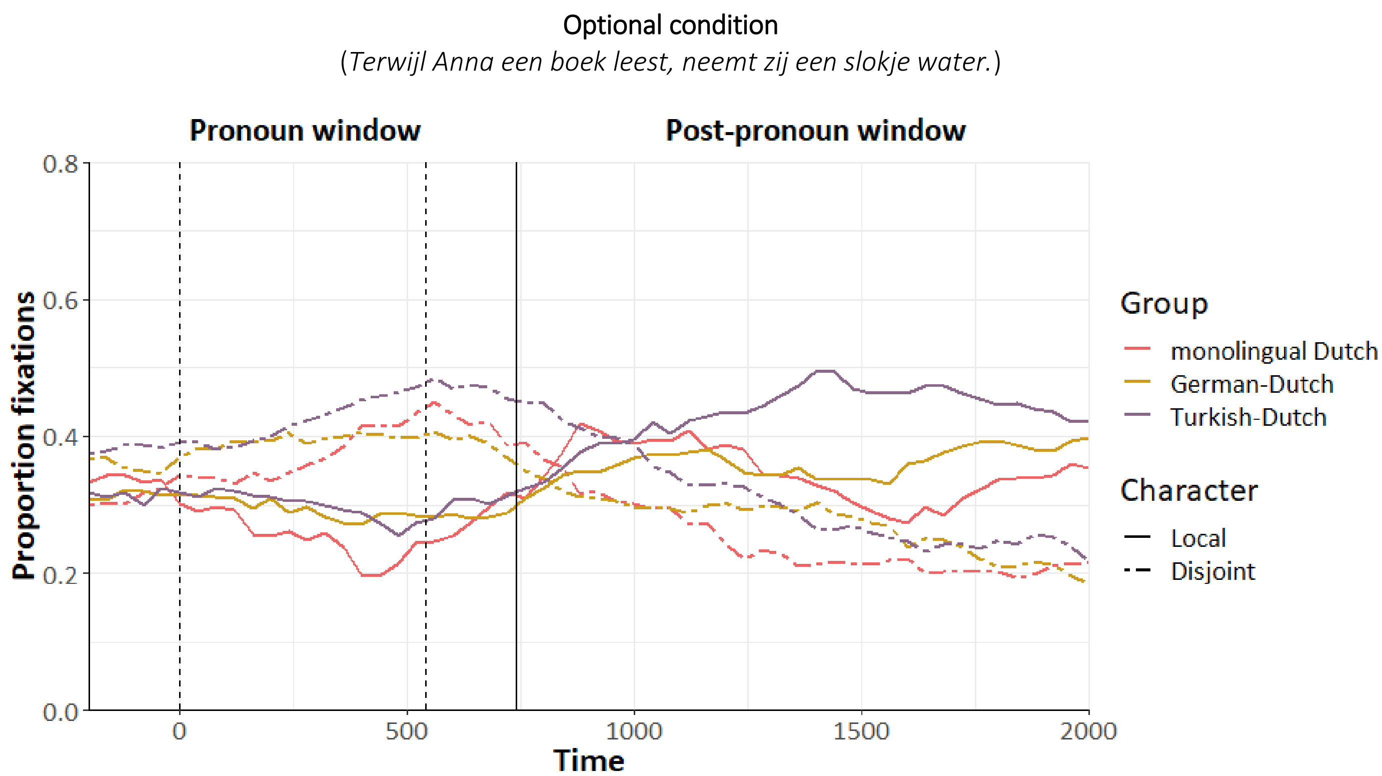


Figure 5.4. Average proportion of fixations on the local and disjoint characters in the optional condition for the three participant groups. The first vertical line indicates the average pronoun onset and the second vertical line indicates the average pronoun offset.

Data analyses

Children's fixations were analysed in R studio (version 4.0.3, R Core Team, 2020), using linear mixed models (lmms) from the lme4 package (version 1.1-23, Bates, Maechler, Bolker, & Walker, 2015) and the lmerTest package (version 3.1-2, Kuznetsova, Brockhoff, & Christensen, 2017). We analysed our predictors of interests, *group* and *time window*, and the interaction between *group* and *time window* for the separate conditions as follows. First, we created a base model including a fixed effect of *trial number*, fixed effects of background variables that significantly predicted children's fixations (*age*, *digit span forward*, *digit span backward*, *Dutch CLT score*), and random intercepts by *participant* and *item*. All continuous fixed effects were centred around their grand mean. We did not include random slopes in the models reported because these typically resulted in convergence errors.

Second, we added the predictors of interest in a stepwise fashion: first, the fixed effect of *group*, then the fixed effect of *time window*, and finally the interaction between *group* and *time window.* Non-significant predictors were kept in the subsequent models. We used Helmert contrasts to test for (i) cross-linguistic influence and (ii) a general bilingualism effect, by comparing: (i) the Turkish-Dutch group (coded as 2/3) to the German-Dutch (-1/3) and the Dutch monolingual group (-1/3); and (ii) the German-Dutch group (-1/2) to the Dutch monolingual group (1/2). In case the latter comparison was significant, we re-ran the model with different Helmert contrasts to explore a potential general bilingualism effect in more detail by comparing: (i) the Turkish-Dutch (-1/3) and the German-Dutch group (-1/3) to the Dutch monolingual group (2/3); and (ii) the Turkish-Dutch group (1/2) to the German-Dutch group (-1/2). Where necessary, the *time window* predictor was re-levelled to investigate effects of *group* within the separate time windows.

Third, to explore significant interactions between *group* and *time window* in more detail, we used generalized additive mixed models (gamms, Wood, 2017; mgcv package: version 1.8-33, Wood, 2020; itsadug package: version 2.4, van Rij, Wieling, Baayen, & van Rijn, 2020; VWPre package: version 1.2.3, Porretta, Kyröläinen, van Rij, & Järvikivi, 2020). The advantage of gamms over lmms is that they can typically handle non-linear patterns as a function of time better than lmms. Furthermore, the mgcv and itsadug packages have built-in functions to deal with autocorrelation over time, a common issue in eye-tracking studies (e.g., Cho, Brown-Schmidt, & Lee, 2018; van Rij, Hendriks, van Rijn, Baayen, & Wood, 2019). Consequently, by means of gamms we could model children's fixations over time by adding so-called smooths of *time* in interaction with *group* (treatment contrasts) to the models. Such smooths allow *time* to follow a non-linear pattern. Model outcomes estimated when during pronoun resolution groups' fixations differed significantly from each other.

Fourth, we tested for the effect of our language dominance factors (*Current input, Cumulative input* and *relative proficiency*) in separate analyses with the data from the Turkish-Dutch children only. In a first step we tested for the main effect of each dominance variable and in a second step we tested for the interactions between *time window* and each dominance variable.

The significance of effects and interactions in all models was tested by comparing the fit of models with and without the effect or interaction of interest using likelihood ratio tests. In all models reported model stress was reduced by removing absolute standardized model residuals above 2.5. Summaries in this section only report effects of our predictors of interest. Complete model summaries can be found in A5.3 in the appendix.

Local condition

Main analyses. Table 5.3 shows the log likelihood tests and summary statistics for the linear mixed models in the local condition. The analyses of children's fixations revealed a non-significant main effect of *group* and a significant main effect of *time window*. Children fixated significantly more on the disjoint than the local referent during the pronoun window and to the local than the disjoint referent during the post-pronoun window. There was also a significant interaction between group and time window. Model summaries showed no significant differences in fixations between the Turkish-Dutch group and the other two groups during either the pronoun or the post-pronoun window, or between the German-Dutch group and the monolingual group during the post-pronoun window. However, during the pronoun window, the German-Dutch children fixated significantly more often on the local compared to the disjoint referent than the monolingual children.

To further explore differences between the groups, we re-ran the model with different Helmert contrasts. The model summaries in Table 5.3 show that during the pronoun window the two bilingual groups fixated significantly more on the local relative to the disjoint referent than the monolingual group and the German-Dutch group significantly more so than the Turkish-Dutch group. Differences between groups were not significant during the post-pronoun window.

Table 5.3. Log likelihood tests and estimates for models (local condition) with main effects of *group* and *time window* and the interaction between group and time window (all children). The reference level column shows which time window was at the reference level. Estimates are given only when the log likelihood test for the corresponding main effect or interaction was significant.

	Log likelihood			Summary statistics					
	X^2	*Δdf*	*p*	Effect	Reference level (window)	*B*	*SE*	*t*	*p*
group	3.0	2	0.228						
time window	2722.0	1	<.001	intercept	pronoun	-0.96	0.132	-7.2	<.001
				intercept	post-pronoun	1.26	0.130	9.6	<.001
				post-pronoun window	pronoun	2.21	0.041	53.5	<.001
group* time window	48.9	2	<.001	Turkish-Dutch vs. others	pronoun	-0.07	0.226	-0.3	.747
				Turkish-Dutch vs. others	post-pronoun	0.00	0.221	0.0	.985
				German-Dutch vs. monolinguals	pronoun	-0.86	0.259	-3.3	.002
				German-Dutch vs. monolinguals	post-pronoun	-0.16	0.254	-0.6	.521
				Bilinguals vs. monolinguals	pronoun	-0.61	0.237	-2.6	.013
				Bilinguals vs. monolinguals	post-pronoun	-0.12	0.232	-0.5	.604
				Turkish-Dutch vs. German-Dutch	pronoun	-0.50	0.245	-2.0	.046
				Turkish-Dutch vs. German-Dutch	post-pronoun	-0.09	0.241	-0.4	.721

We further explored the groups' behaviour over time using gamms. The smooth of time by group significantly improved the model fit (X^2 = 10.7; Δdf = 6; p = .002). This suggested that the development of children's fixation patterns differed over time between groups. Visualizations of the smooth (see A5.4 in the appendix) showed that the German-Dutch children fixated significantly more on the local relative to the disjoint referent compared to the Turkish-Dutch bilingual children during both the pronoun and the post-pronoun window (0 ms – 1010 ms and 1657 ms – 2000 ms) and compared to the monolingual children for almost the entire pronoun window (20 ms – 323 ms and 384 ms – 707 ms). Furthermore, the monolingual children fixated significantly more on the local relative to the disjoint referent than the Turkish-Dutch children at the end of the post-pronoun window (1919 ms – 2000 ms).

Language dominance analyses. The relationship between *relative proficiency* and fixations on the local and disjoint referent is plotted in Figure 5.5. Figures for *current* and *cumulative input* can be found in the appendix (A5.5). Log likelihood tests and estimates of separate models with our three dominance measures are shown in Table 5.4. The main effect of *relative proficiency* approached significance. The higher children's CLT score in Turkish relative to Dutch, the less they looked at the disjoint relative to the local referent. The main effect of *relative proficiency* was significantly moderated by *time window*. Summaries showed that the observed effect of *relative proficiency* was significant only in the pronoun window.

The interactions between *current input* and *time window* and *cumulative input* and *time window* were significant as well. Summaries showed that the effect of *current input* was larger in the pronoun window than in the post-pronoun window, similar to *relative proficiency*. In contrast, the effect of *cumulative input* was larger in the post-pronoun window than in the pronoun window. Simple effects of *current* and *cumulative input* were not significant in either time window, however.

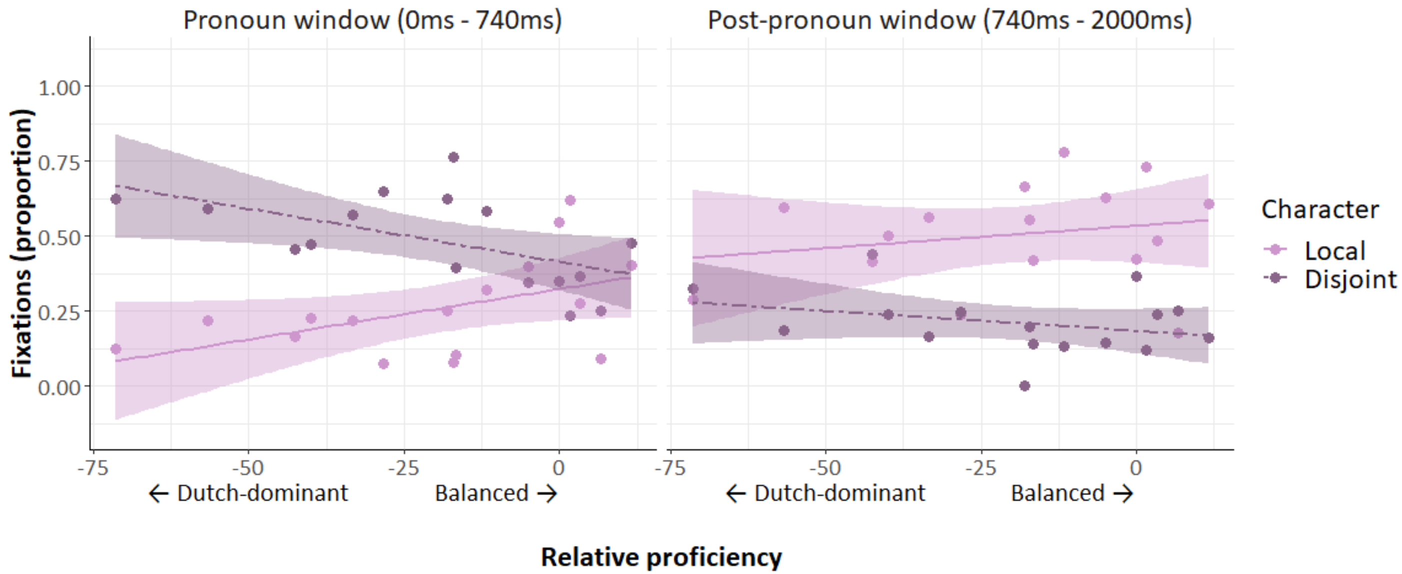


Figure 5.5. Turkish-Dutch children's average proportions of fixations on the local and disjoint referent in the local condition by time window by relative proficiency. The shaded areas display the 95% confidence intervals.

Table 5.4. Log likelihood tests and estimates for models (local condition) with main effects of the dominance measures and interactions between the dominance measures and *time window* (Turkish-Dutch children only). The reference level column shows which time window was at the reference level. Estimates are given only when the log likelihood test for the corresponding main effect or interaction was significant.

	Log likelihood			Summary statistics					
	X^2	*Δdf*	*p*	Effect	Reference level (window)	*B*	*SE*	*t*	*p*
relative proficiency	3.5	1	.060	relative proficiency	n.a.	-0.02	0.010	-1.9	.079
****time window***	33.7	1	<.001	relative proficiency	pronoun	-0.03	0.010	-3.3	.005
				relative proficiency	post-pronoun	-0.01	0.010	-1.1	.274
				relative proficiency *time window	pronoun	0.02	0.004	5.8	<.001
current input	0.6	1	.442						
****time window***	7.3	1	.007	current input	pronoun	0.04	0.028	1.3	.227
				current input	post-pronoun	0.01	0.279	0.4	.722
				current input *time window	pronoun	-0.03	0.009	-2.7	.007
cumulative input	0.8	1	.381						
****time window***	45.6	1	<.001	cumulative input	pronoun	-0.01	0.018	-0.7	.488
				cumulative input	post-pronoun	0.03	0.018	1.8	.088
				cumulative input *time window	pronoun	0.05	0.007	6.8	<.001

Disjoint condition

Main analyses. Table 5.5 shows the log likelihood tests and summary statistics for the linear mixed models in the disjoint condition. The analyses of children's fixations revealed a non-significant main effect of *group* and a significant main effect of *time window*. Summaries showed that children fixated significantly more on the disjoint than the local referent during the pronoun window and during the post-pronoun window. The difference in fixations was significantly larger in the post-pronoun window. There was also a significant interaction between group and time window. Summaries showed that the Turkish-Dutch children looked more to the disjoint relative to the local referent during the post-pronoun window compared to the other two groups. This difference approached significance. None of the other group comparisons approached significance.

We further explored the groups' behaviour over time using gamms. The smooth of time by group significantly improved the model fit (X^2 = 34.6; Δdf = 6; p = < .001). This showed that the development of children's fixation patterns differed over time between groups. Visualizations of the smooth, however, revealed no significant differences between the groups over time (see A5.4 in the appendix for the figures).

Table 5.5. Log likelihood tests and estimates for models (disjoint condition) with main effects of *group* and *time window* and the interaction between group and time window (all children). The reference level column shows which time window was at the reference level. Estimates are given only when the log likelihood test for the corresponding main effect or interaction was significant.

	Log likelihood			**Summary statistics**					
	X^2	*Δdf*	*p*	**Effect**	**Reference level (window)**	*B*	*SE*	*t*	*p*
group	2.1	2	.344						
time window	29.9	1	<.001	intercept	pronoun	-0.74	0.181	-4.4	<.001
				intercept	post-pronoun	-0.98	0.180	-5.4	<.001
				post-pronoun window	pronoun	-0.24	0.043	5.5	<.001
group* time window	31.7	2	<.001	Turkish-Dutch vs. others	pronoun	-0.09	0.329	-0.3	.796
				Turkish-Dutch vs. others	post-pronoun	-0.59	0.325	-1.8	.074
				German-Dutch vs. monolinguals	pronoun	0.14	0.378	0.4	.715
				German-Dutch vs. monolinguals	post-pronoun	0.42	0.374	1.1	.268

Language dominance analyses. The relationship between *relative proficiency* and fixations on the local and disjoint referent is plotted in Figure 5.6. Figures for *current* and *cumulative input* can be found in the appendix (A5.5). Log likelihood tests and estimates of separate models with our three dominance measures are shown in Table 5.6. The only significant main effect was observed for *relative proficiency*. The higher children's CLT score in Turkish relative to Dutch, the less they looked at the disjoint relative to the local referent. The main effect of *relative proficiency* was not significantly moderated by *time window*. This shows that the effect of *relative proficiency* was similar during the pronoun and the post-pronoun windows.

The interactions between *current input* and *time window* and between *cumulative input* and *time window* were significant. Summaries showed that the effect of *current input* was significantly larger in the post-pronoun window than in the pronoun window. Only in the post-pronoun window did the simple effect of *current input* approach significance. The direction of the estimate showed that the more children were exposed to Turkish relative to Dutch at the time of testing, the less they fixated on the disjoint compared to the local referent. In contrast, the effect of *cumulative input* was significantly stronger in the pronoun window than in the post-pronoun window. There was a non-significant trend in the pronoun window: the more cumulative exposure children had received in Turkish compared to Dutch, the less they looked at the disjoint referent relative to the local referent.

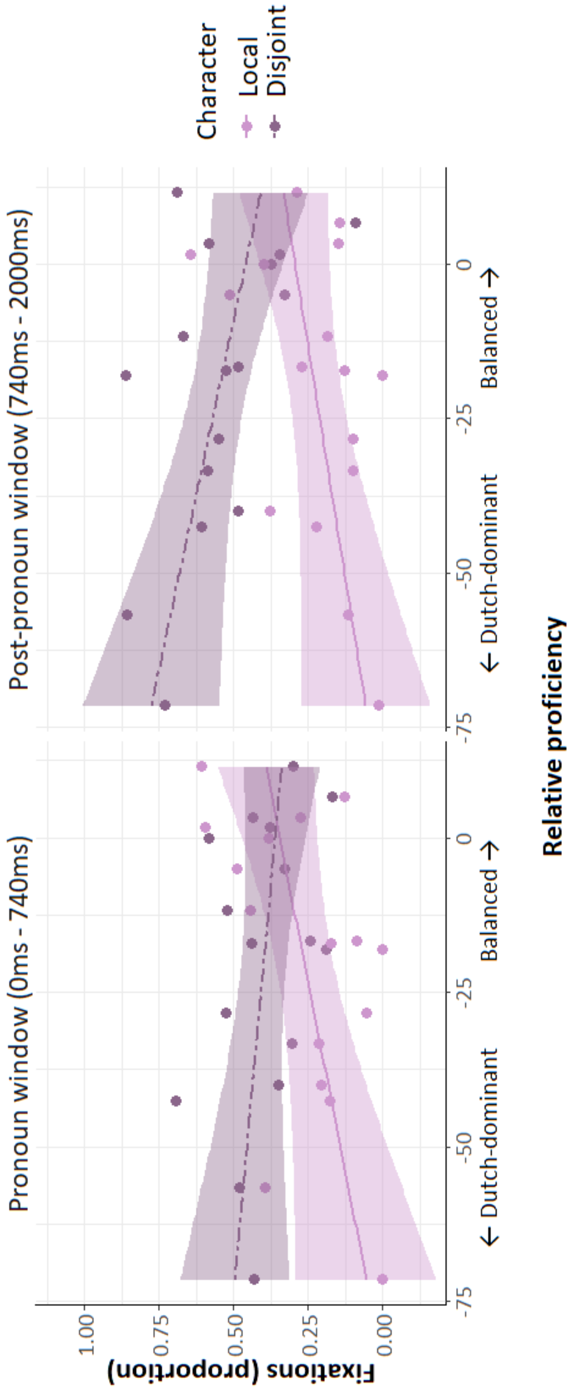


Figure 5.6. Turkish-Dutch children's average proportions of fixations on the local and disjoint referent in the disjoint condition by time window by relative proficiency. The shaded areas display the 95% confidence intervals.

Table 5.6. Log likelihood tests and estimates for models (disjoint condition) with main effects of the dominance measures and interactions between the dominance measures and *time window* (Turkish-Dutch children only). The reference level column shows which time window was at the reference level. Estimates are given only when the log likelihood test for the corresponding main effect or interaction was significant.

	Log likelihood			Summary statistics					
	X^2	*Δdf*	*p*	Effect	Reference level (window)	*B*	*SE*	*t*	*p*
relative proficiency	6.8	1	.009	relative proficiency	n.a.	-0.31	0.115	-2.7	.016
****time window***	1.6	1	.205						
current input	2.7	1	.103						
****time window***	19.9	1	<.001	current input	pronoun	0.03	0.035	0.7	.478
				current input	post-pronoun	0.07	0.035	2.0	.067
				current input *time window	pronoun	0.04	0.010	-4.4	<.001
cumulative input	0.7	1	.408						
****time window***	27.4	1	<.001	cumulative input	pronoun	0.43	0.247	1.7	.103
				cumulative input	post-pronoun	0.06	0.243	0.2	.822
				cumulative input *time window	pronoun	-0.37	0.071	-5.2	<.001

Optional condition

Main analyses. Table 5.7 shows the log likelihood tests and summary statistics for the linear mixed models in the optional condition. The analyses of children's fixations revealed a non-significant main effect of *group* and a significant main effect of *time window*. Summaries of the model showed that children fixated significantly more on the disjoint than the local referent during the pronoun window and to the local than the disjoint referent during the post-pronoun window. This difference between time windows was significant. There was also a significant interaction between group and time window. Summaries, however, showed no significant differences in fixations between the Turkish-Dutch group and the other two groups, or between the German-Dutch group and the monolingual group.

We further explored the groups' behaviour over time using gamms. The smooth of time by group did not significantly improve the model fit (X^2 = 4.7; Δdf = 6; p = .153). Thus, the development of children's fixation patterns did not differ over time between groups.

Language dominance analyses. The relationship between *relative proficiency* and fixations on the local and disjoint referent is plotted in Figure 5.7. Figures for *current* and *cumulative input* can be found in the appendix (A5.5). Log likelihood tests and estimates of separate models with our three dominance measures are shown in Table 5.8. The main effect of *relative proficiency* was significant. The higher children's CLT score in Turkish relative to Dutch, the less they fixated on the disjoint relative to the local referent. The main effect of *relative proficiency* was not significantly moderated by *time window*. This shows that the effect of *relative proficiency* was similar during the pronoun and the post-pronoun windows. There were no significant main effects of or interactions with *current input* and *cumulative input*.

Table 5.7. Log likelihood tests and estimates for models (optional condition) with main effects of *group* and *time window* and the interaction between group and time window (all children). The reference level column shows which time window was at the reference level. Estimates are given only when the log likelihood test for the corresponding main effect or interaction was significant.

	Log likelihood			Summary statistics					
	X^2	*Δdf*	*p*	Effect	Reference level (window)	*B*	*SE*	*t*	*p*
group	0.1	2	.964						
time window	547	1	<.001	intercept	pronoun	0.55	0.169	-3.3	.002
				intercept	post-pronoun	0.45	0.168	2.7	.009
				post-pronoun window	pronoun	1.00	0.043	23.5	<.001
group* time window	11.6	2	.003	Turkish-Dutch vs. others	pronoun	-0.26	0.319	-0.8	.429
				Turkish-Dutch vs. others	post-pronoun	0.00	0.316	0.0	.989
				German-Dutch vs. monolinguals	pronoun	-0.09	0.367	-0.3	.800
				German-Dutch vs. monolinguals	post-pronoun	0.07	0.363	0.2	.849

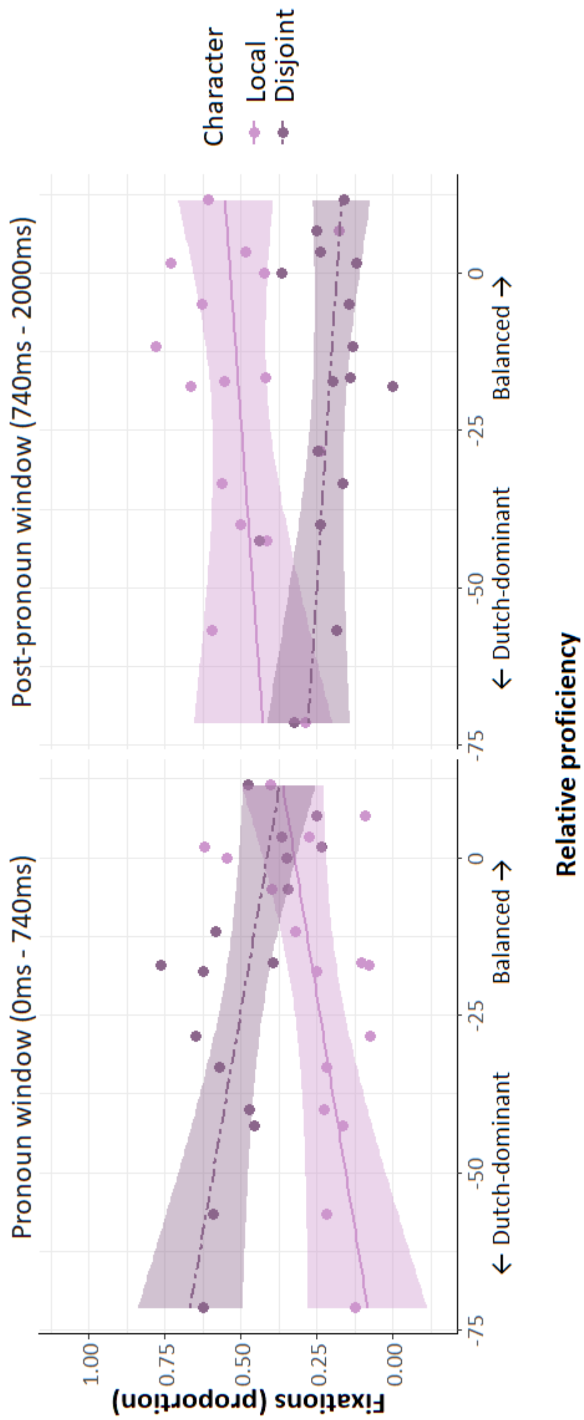


Figure 5.7. Turkish-Dutch children's average proportions of fixations on the local and disjoint referent in the optional condition by time window by relative proficiency. The shaded areas display the 95% confidence intervals.

Table 5.8. Log likelihood tests and estimates for models (disjoint condition) with main effects of the dominance measures and interactions between the dominance measures and *time window* (Turkish-Dutch children only). The reference level column shows which time window was at the reference level. Estimates are given only when the log likelihood test for the corresponding main effect or interaction was significant.

	Log likelihood			Summary statistics					
	X^2	*Δdf*	*p*	Effect	Reference level (window)	*B*	*SE*	*t*	*p*
relative proficiency	5.1	1	.023	relative proficiency	n.a.	0.03	0.013	2.3	.035
****time window***	1.1	1	.284						
current input	1.0	1	.324						
****time window***	0.1	1	.735						
cumulative input	1.0	1	.320						
****time window***	1.0	1	.315						

To summarize, in all three conditions, children fixated significantly more on the disjoint referent compared to the local referent during the pronoun window. In the local and optional condition, this effect reversed during the post-pronoun window. In contrast, in the disjoint condition, children continued to fixate more on the disjoint relative to the local referent during the post-pronoun window. This effect was stronger than during the pronoun window.

In all three conditions, the effect of *time window* was modified by *group*. In the local condition, the disjoint preference in the pronoun window was smaller in the bilingual groups compared to the monolingual group and smaller in the German-Dutch group than in the Turkish-Dutch group. In the more detailed time analyses (gamms), we observed similar differences between the German-Dutch, but not the Turkish-Dutch and monolingual children. Instead, for a short period of time (< 100 ms) at the end of the post-pronoun window the monolingual children had a stronger local referent preference than the Turkish-Dutch children. Furthermore, during the pronoun and the post-pronoun windows, the German-Dutch children also fixated less on the disjoint relative to the local referent than the Turkish-Dutch children. In the disjoint condition, there was a non-significant trend in the analyses collapsed over time for a stronger preference for the disjoint referent in the Turkish-Dutch children than the other two groups in the post-pronoun window. This effect was no longer visible in the analyses over time, however. In the optional condition, there were no significant differences between groups in the analyses collapsed over time, nor in the time course analyses.

Finally, Turkish-Dutch children's language dominance profiles affected their fixation patterns in all three conditions. This effect was most pronounced for the *relative proficiency* measure. To be more precise, the higher children's Turkish CLT score relative to their Dutch CLT score, the less they fixated on the disjoint referent compared to the local referent. This pattern was significant in both time windows in all conditions except in the post-pronoun window in the local condition. Similar though non-significant trends could be observed in the pronoun window for *current input* in the local condition and *cumulative input* in the disjoint condition and in the post-pronoun window for *cumulative input* in the local condition and for *current input* in the disjoint condition.

Offline pronoun interpretation

Figure 5.8 shows the proportion of children's responses on the Dutch pronoun task where the local referent was selected in the different groups (see A5.6 in the appendix for the groups' mean percentage of local pronoun choices).

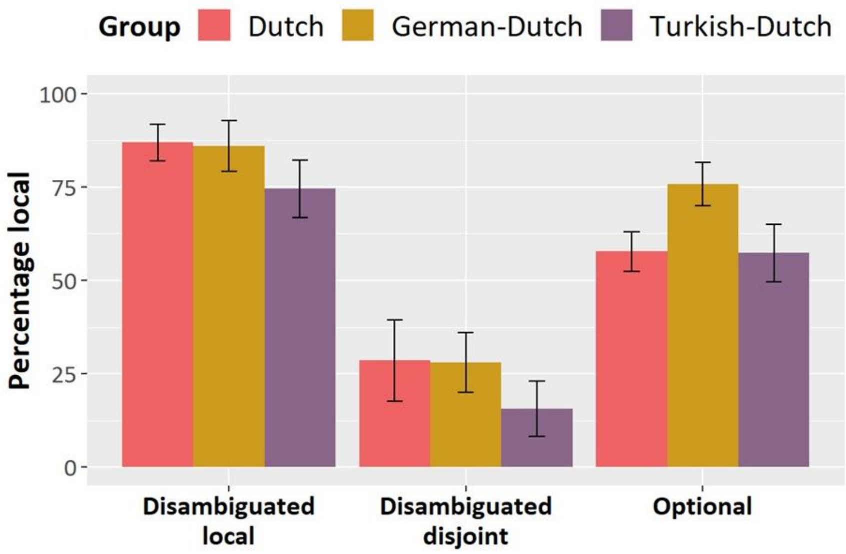


Figure 5.8. Average percentage of choices for the local referent on the Dutch pronoun task by *group* and *condition*. Error bars represent standard errors.

Data analyses

Children's local choices were analysed using generalized lmms with children's pronoun choices as binary dependent variables (local = 1, disjoint = 0). First, we created a base model including a fixed effect of *trial number*, fixed effects of background variables that significantly predicted children's fixations, and random intercepts by *participant*. Random intercepts by *Item* and random slopes were dropped from the models as these typically resulted in convergence errors. All continuous fixed effects were centred around their grand mean.

Second, we added the fixed effect of *condition*, then the fixed effect of *group*, and finally the interaction between *condition* and *group.* Contrasts for *condition* and *group* were Helmert coded. For *condition* we compared children's local referent choices in the disjoint condition (2/3) to the local (-1/3) and optional condition (-1/3) combined, and in the local condition (1/2) compared to the optional condition (-1/2). For *group* we compared local referent choices in the Turkish-Dutch group (2/3) to the German-Dutch (-1/3) and monolingual group (-1/3) combined, and in the German-Dutch (1/2) compared to the Dutch monolingual group (-1/2). To explore significant interactions between *condition* and *group,* we used treatment contrasts for the different conditions and explored effects of group using Helmert contrasts by relevelling conditions.

Third, we tested for the effect of our language dominance factors (*Current input, Cumulative input* and *CLT difference score*) in separate analyses with the data from the Turkish-Dutch children only. In a first step, we tested for the main effect of each dominance variable and in a second step we tested for the interaction between *condition* and each dominance variable. Again, conditions were treatment coded and relevelled where necessary.

The significance of effects and interactions in all models was tested by comparing the fit of models with and without the effect or interaction of interest using likelihood ratio tests. Where possible, model stress was reduced by removing absolute model residuals above 2.5. Summaries with significant main effects of and interactions with our variables of interest can be found in A5.7 in the appendix.

Main analyses

Table 5.9 shows the log likelihood tests and summary statistics for the generalized lmms. There was a significant main effect of condition, but no significant main effect of group. The model summary with the main effect of *condition* showed that children were significantly less likely to choose the local referent in the disjoint condition than in the other two conditions and in the optional condition than in the local condition. There was also a significant interaction between group and condition. Turkish-Dutch children were less likely to choose the local referent than the other two groups in each condition. However, these differences were not significant. Furthermore, there were no significant differences between the German-Dutch and the monolingual children in any condition. Hence, the significant interaction between group and condition was not driven by any of the comparisons of interest.

Table 5.9. Log likelihood tests and estimates for models of children' s offline pronoun choices with main effects of *condition* and *group* and the interaction between *group* and *condition* (all children) The reference level column shows which condition was at the reference level. Estimates are given only when the log likelihood test for the corresponding main effect or interaction was significant.

	Log likelihood			Summary statistics					
	X^2	*Δdf*	*p*	Effect	Reference level (condition)	*B*	*SE*	*z*	*p*
condition	909.5	2	<.001	disjoint vs. other	n.a.	-5.63	0.334	-16.8	<.001
				local vs. optional	n.a.	-1.71	0.201	8.5	<.001
group	2.8	2	.250						
group* condition	22.0	4	<.001	Turkish-Dutch vs. others	local	-1.17	0.726	-1.6	.106
				Turkish-Dutch vs. others	disjoint	-1.16	0.742	-1.6	.118
				Turkish-Dutch vs. others	optional	-0.73	0.706	-1.0	.304
				German-Dutch vs. monolinguals	local	0.71	0.861	0.8	.409
				German-Dutch vs. monolinguals	disjoint	-0.46	0.825	-0.6	.574
				German-Dutch vs. monolinguals	optional	1.53	0.821	1.9	.062

Language dominance

The relationship between *relative proficiency* and children's offline referent choices is plotted in Figure 5.9. Figures for *current* and *cumulative input* can be found in the appendix (A5.8). Table 5.10 shows the effects of our three dominance variables.

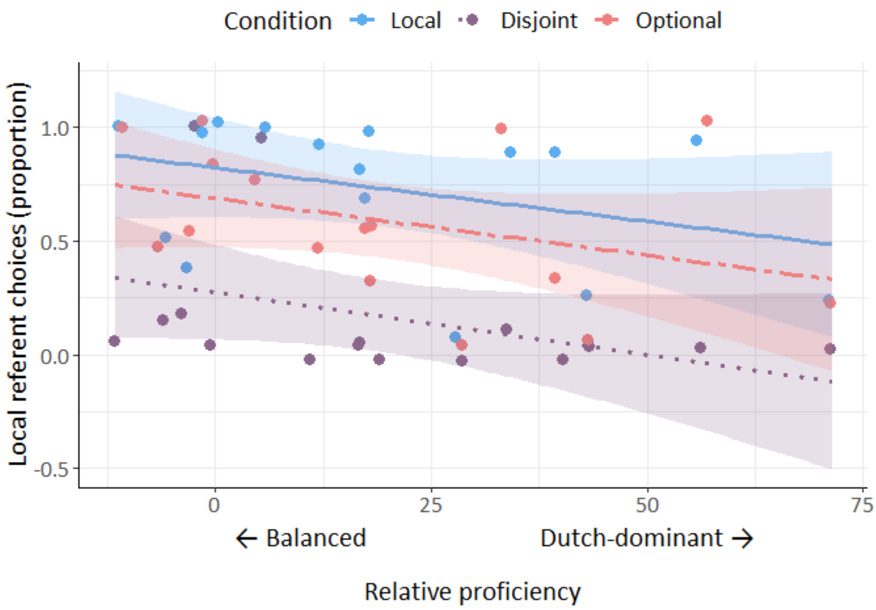


Figure 5.9. Turkish-Dutch children's average proportion of offline local referent choices by *condition* by their *relative proficiency* score. The shaded areas display the 95% confidence intervals. The two outliers discussed in the text are the purple datapoints at the top-right. Note: some datapoints seem to have a negative value or a value larger than 1. This was not the case. In the figure, we allowed some deviation in the vertical and horizontal position of the datapoints, because otherwise some points would overlap completely.

Table 5.10. Log likelihood tests and estimates for models of children's offline pronoun choices with main effects of the dominance measures and interactions between the dominance measures and *condition* (Turkish-Dutch children only). The reference level column shows which condition was at the reference level. Estimates are given only when the log likelihood test for the corresponding main effect or interaction was significant.

	Log likelihood			Summary statistics					
	X^2	*Δdf*	*p*	Effect	Reference level (condition)	*B*	*SE*	*z*	*p*
relative proficiency	3.1	1	.076	relative proficiency	n.a.	-0.05	0.029	-1.8	.066
****condition***	7.7	2	.021	relative proficiency	local	-0.03	0.023	-1.4	.149
				relative proficiency	disjoint	-0.09	0.029	-3.0	.003
				relative proficiency	optional	-0.04	0.023	-1.5	.125
current input	0.1	1	.773						
****condition***	0.8	2	.676						
cumulative input	0.3	1	.581						
****condition***	1.4	2	.504						

There is a marginal significant main effect of *relative proficiency*. The summary shows that the better children's Turkish CLT score compared to their Dutch CLT score, the more likely they were to choose the local referent. This effect was significantly modified by *condition*. Whilst the direction of the effect of *relative proficiency* was similar in all three conditions, it only reached significance in the disjoint condition. However, closer inspection of the individual children's data suggested that the effect in the disjoint condition was carried by two children with a relatively high *relative proficiency score* who (almost) always chose the local referent. After removal of these children's data, the main effect of *relative proficiency* still approached significance (X^2 = 3.4; Δdf = 1; p = .064). However, the interaction between *relative proficiency* and *condition* was no longer significant (X^2 = 3.7; Δdf = 2; p = .154). There were no significant effects of or interactions with the other language dominance measures.

To summarize, children chose the local referent most often in the local condition, least often in the disjoint condition, with the optional condition falling somewhere in-between. The Turkish-Dutch children chose the local referent less often when compared to the other two groups. However, this effect failed to reach significance in any condition. There was a non-significant trend for *relative proficiency*: the better children's CLT score in Turkish compared to Dutch, the more likely they were to choose the local referent. This effect reached significance only in the disjoint condition. However, upon closer inspection of the data the effect in the disjoint condition appeared to be mainly carried by the local responses of two children. After removal of their data, only the general trend remained irrespective of condition.

5.5 Discussion

Before we turn to the discussion of the findings in relation to our hypotheses, we first briefly focus on the general patterns found in children's fixations during pronoun processing. Children showed an initial preference for the disjoint referent over the local referent when listening to the pronouns. This was unexpected, as previous eye-tracking studies with monolingual children and L2 learners either showed no initial preference for a topic antecedent (e.g., Arnold et al., 2007; Cunnings et al., 2017; Song & Fisher, 2005) or an early preference for the topic antecedent (e.g., Contemori & Dussias, 2020) in non-null subject languages.

There are at least two explanations for our different findings. First, the use of *terwijl* ("while") might have signalled a shift in topic in the main clause, resulting in an initial disjoint preference. This would be similar to observations made for adults processing clauses in French containing *avant* ("before"; e.g., Hemforth et al., 2010). We are not aware on any processing studies on the function of *terwijl* ("while") in Dutch. Second, children might have had an initial expectation for both characters in each story to perform an action, rather than just one character. In particular, children might have first considered the disjoint character to perform the action in the main clause, after having listened to the local character performing an action in the *while*-clause. This would explain why children shifted their attention to the disjoint character in the main clause.

Note that the task faced by children in our experiment was to suppress their initial disjoint fixations in the local and optional conditions in favour of the local referent. To be more precise, they had to reanalyse their initial disjoint interpretation in these conditions. Our online and offline results showed that overall, children were successful at this: (i) they fixated significantly more on the local referent after having heard the pronoun in the local and optional condition; and (ii) they chose the local referent significantly more often in the local and optional conditions than in the disjoint condition. In other words, children's behaviour showed that they were able to make use of gender and discourse information in their online and offline pronoun resolution.

We now turn to our findings and discuss them in relation to our hypotheses. Subsequently, we propose an account in terms of processing to explain the observed behaviour in the Turkish-Dutch children.

Online group comparisons

In our group comparisons, there were a number of differences between the Turkish-Dutch children and the other two participant groups. In particular, there was some indication that the Turkish-Dutch children fixated more on the disjoint referent than the German-Dutch and the monolingual group, namely, in the local and disjoint conditions. This was in line with our first hypothesis that the discourse-pragmatic status of the Turkish overt pronoun results in more disjoint fixations during overt pronoun resolution in Dutch. Similar results were obtained by Schimke et al. (2018) for adult L2 learners of Spanish.

However, the larger number of disjoint fixations in the Turkish-Dutch group compared to the other two groups was not stable. In the first place, in the local condition the difference between the Turkish-Dutch and monolingual children emerged only at the end of the time window investigated and for a very short amount of time (< 100 ms). In the second place, the differences in the disjoint condition were only marginally significant in the analyses collapsed over time and disappeared completely when the time course of fixations was taken into account. Furthermore, there were no differences between the groups in the optional condition. Hence, evidence for cross-linguistic influence in the group analyses was only weak, at best. Therefore, our data did not convincingly support the first hypothesis that Turkish-Dutch children fixate more on the disjoint referent under influence of their experience with Turkish.

Our findings are in line with studies on the use and offline interpretation of pronouns by bilingual children that found no evidence for cross-linguistic influence from a null subject language into a non-null subject language (e.g., Argyri & Sorace, 2007; Schmitz et al., 2011; Serratrice et al., 2004; Sorace et al., 2009). Furthermore, our observations are also consistent with online studies on pronoun resolution in L2 learners that found no evidence for cross-linguistic influence from a null subject language into a non-null subject language (e.g., Contemori & Dussias, 2020; Cunnings et al., 2017), and, in particular, from Turkish into Dutch (Roberts et al., 2008).

The role of language dominance

In our analyses of language dominance, we observed that the more balanced children were in Turkish and Dutch, as opposed to being dominant in Dutch, the less they fixated on the disjoint referent. This relation was most pronounced when language dominance was operationalised using relative proficiency: the more proficient children were in Turkish relative to Dutch – as measured by a lexical proficiency task – the less they fixated on the disjoint referent in all conditions and all time windows tested, except for the post-pronoun window in the local condition. Similar, but non-significant, trends were observed for our other two dominance measures, current and cumulative input: the more children were exposed to Turkish at the time of testing and the more they had been exposed over the years compared to Dutch, the less they fixated on the disjoint referent relative to the local referent.

These effects of language dominance corroborated the second hypothesis, but only partially. We predicted effects of Turkish on Dutch pronoun resolution to become stronger with increased dominance in Turkish. This was indeed what happened. Our study is, to the best of our knowledge, the first to find evidence of cross-linguistic influence in bilingual children's pronoun resolution from a null subject language into a non-null subject language (e.g., Argyri & Sorace, 2007; Sorace et al., 2009). The effects of cross-linguistic influence became visible when we took language dominance into account. This is in line with previous offline studies on bilingual children's pronoun resolution in a null subject language (e.g., Argyri & Sorace, 2007; Sorace et al., 2009) and on other morphosyntactic properties (e.g., van Dijk, van Wonderen, et al., Chapter 2).

Our observation that online effects of dominance were most pronounced for children's relative language proficiency and less for current and cumulative language exposure is in accordance with an online study by van Dijk, Dijkstra and Unsworth (Chapter 3). This self-paced listening study tested whether current and cumulative input and relative proficiency – as measured by a sentence repetition test – moderated online cross-linguistic influence in simultaneous bilingual children. In this study we observed that such effects were most pronounced for relative proficiency and were limited or insignificant for the input measures, in line with our present study. This similarity between studies suggests that before cross-linguistic influence can take place during sentence processing, morphosyntactic knowledge first has to be acquired and stable connections formed between knowledge representations within and between languages (e.g., Hartsuiker & Bernolet, 2017). Language proficiency might be a better measure of the organization of such knowledge representations than current and cumulative input, especially in children who are still in the process of acquiring their languages. It is as yet an open question whether or not it is the relative proficiency between languages that drives the effect of dominance or the absolute proficiency in the languages. In order to answer this question, a measure of language dominance is necessary that takes into account both absolute proficiency and relative proficiency in bilingual children's languages (e.g., Blom, Küntay, Messer, Verhagen, & Leseman, 2014; Goriot, Broersma, McQueen, Unsworth, & van Hout, 2018).

The direction of cross-linguistic influence went in the opposite direction than predicted. We expected *more* disjoint fixations with increased dominance in Turkish, in line with a 'Turkish discourse strategy'. Instead, we observed *fewer* disjoint fixations. The direction of the effect contrasts with the behaviour of the L1 Spanish speakers in the eye-tracking task by Schimke and colleagues (2018), who fixated *more* on the disjoint referent during the eye-tracking task than the control group of L1 French speakers, that is, in line with pronoun preferences in Spanish. In our view, the different outcomes between studies can be accounted for by differences in the dominance profiles of the participants studied. The participants in Schimke et al. (2018) were tested in their L2, which is likely to be their non-dominant language. Participants in our study were either dominant in the language tested or balanced in their languages. We will elaborate on the direction of the effect and the role of language dominance in online sentence processing in the sections below.

General effect of bilingualism

To test for a general effect of bilingualism, we also compared the behaviour of the two bilingual groups together with that of the monolingual group. In the local condition there was some evidence that the two bilingual groups initially focused more on the local referent than the monolingual children, in line with our hypothesis. However, analyses over time showed that this effect was caused by the German-Dutch children only. Furthermore, in the other two conditions we found no evidence for the two bilingual groups behaving similarly to each other but differently from the monolingual group. Therefore, we reject the third hypothesis that online processing in bilingual children is less efficient than in monolingual children due to a general effect of bilingualism.

Our observations differ from previous online pronoun studies with L2 adults (e.g., Cunnings et al., 2017; Roberts et al., 2008) and Sorace et al.'s (2009) offline pronoun study with bilingual children, where a general bilingualism effect was attested. Our results show that observed difficulties during sentence processing as observed in adult L2 learners due to bilingualism (e.g., Clahsen & Felser, 2006; Hopp, 2010; Sorace, 2011) should not automatically be extended to simultaneous bilinguals (also see Felser, 2020 and van Dijk, Dijkstra & Unsworth, Chapter 4).

Offline versus online cross-linguistic influence

Finally, the picture-selection data patterned similarly to the online results. First, the Turkish-Dutch children chose the disjoint referent slightly – although not significantly – more often than the other two groups in the local and disjoint condition. Second, the more balanced children were in Turkish and Dutch, as opposed to being dominant in Dutch, the more often they chose the local referent as pronoun antecedent. Nevertheless, whilst the effects of language dominance were statistically significant online, they failed to reach significance offline. The only exception was the disjoint condition: in this condition language dominance significantly predicted children's offline referent choices. However, this effect was carried by two children only and once they were removed, the effect no longer approached significance.

Our findings support our fourth hypothesis that online and offline effects of cross-linguistic influence pattern similarly, but are more pronounced online. We consider two explanations for these online-offline differences. We first focus on differences between the constructs which online and offline tasks are each understood to measure. We then discuss the relationship between offline and online data in the context of a more general account of our findings in the next section.

Online measures are considered more direct measures of language processing than offline measures (e.g., Marinis, 2010). Offline tasks do not only tap into children's linguistic knowledge, but also allow for more explicit strategies and use of metalinguistic abilities. Furthermore, children's responses are measured *after* a sentence has been processed. This means that children have to process a sentence, keep the information in working memory and make a decision. This poses a demand on working memory. Consequently, children's offline referent-choices in our study might contain artefacts of memory capacity limitations and strategies that obscure effects of cross-linguistic influence and language dominance. This noise in the offline data might have impacted the power of the analyses (e.g., Brysbaert, 2019) and, consequently, decreased the chances of obtaining a significant effect. Given that the number of children tested in this study was relatively small (due to the consequences of Covid-19), the analyses of the offline data will likely have lacked the power to detect significant effects, if they were there. This is a common issue in offline comprehension and production studies on cross-linguistic influence in bilingual children (see van Dijk, van Wonderen et al., Chapter 2). Instead, the clear and significant dominance effects in the online data suggest that our eye-tracking technique was relatively robust against small sample sizes. Hence, our findings imply that online techniques might be better suited to detecting (subtle) effects of cross-linguistic influence in bilingual children (e.g., van Dijk, Dijkstra & Unsworth, Chapter 3; see Martohardjono, Phillips, Madsen II, & Schwartz, 2017, for a similar observation for early bilingual adults).

In sum, we found evidence for cross-linguistic influence during online pronoun resolution from Turkish into Dutch when we took children's language dominance profiles into account. The effect of cross-linguistic influence went in the opposite direction than predicted, however. In particular, the more balanced bilingual children were in their languages, as opposed to being Dutch-dominant, the less Turkish-like they behaved. In the following section, we show how this at first sight unexpected finding follows from an account of cross-linguistic influence based on language co-activation and inhibition.

Mechanisms underlying cross-linguistic influence

There is ample evidence that bilinguals' lexical and syntactic representations in one language become activated and primed for subsequent use whilst they are processing the other language (e.g., Dijkstra & van Heuven, 2002; Hartsuiker, Pickering, & Veltkamp, 2004; Loebell & Bock, 2003; Marian & Spivey, 2003). It has even been argued that such co-activation and priming over time results in shared syntactic representations between languages (e.g., Hartsuiker et al., 2004; Hartsuiker & Bernolet, 2017). Most of this work is with bilingual adults. Still, a number of recent studies show similar effects in bilingual children (e.g., Hsin, Legendre, & Omaki, 2013; Vasilyeva et al., 2010; Von Holzen & Mani, 2012). In other words, lexical and syntactic representations from both bilingual children's languages compete for selection during language processing (e.g., Nicoladis, 2006, 2012; Nicoladis et al., 2010). As a consequence, in order to correctly interpret sentences bilingual children have to inhibit representations from the language not in use and select representations from the language being processed (e.g., Green, 1998; Green & Abutalebi, 2013; Meuter & Allport, 1999).

Lemmas play a central role in accounts of cross-language co-activation (e.g., Dijkstra & van Heuven, 2002; Green, 1998; Hartsuiker et al., 2004). Lemmas are lexical representations that serve as an interface between the form, meaning and morphosyntactic information of words (e.g., Levelt et al., 1999). For example, in models of syntactic priming lemmas are stored together with information about which sentence structures they can appear in, represented by so-called combinatorial nodes (e.g., Pickering & Branigan, 1998). Importantly, lemmas from one language can activate lemmas from the other language in bilinguals, either through shared semantic representations or through shared combinatorial nodes (e.g., Dijkstra & van Heuven, 2002; Hartsuiker et al., 2004).

On such an approach, the processing of overt pronouns in Dutch by Turkish-Dutch children activates the form of the pronoun (e.g., *zij*), the lemma for Dutch pronouns and its semantic representations (i.e., a topic and a non-topic interpretation). In turn, the Turkish pronoun lemma and it forms become activated through spreading activation from the semantic representations shared with Dutch. This is schematically depicted in Figure 5.10 for the Dutch pronoun *zij* (based on Hartsuiker, Pickering and Veltkamp's 2004 model of structural priming in L2 learners).[2] However, the Dutch and Turkish pronoun lemmas have different preferences: whilst the Dutch overt

[2] Because the processing of overt pronouns activates discourse-pragmatic principles relevant to overt pronoun resolution (e.g., Serratrice, 2007), we assume that the processing of overt pronouns in Dutch activates only overt pronouns and their interpretations in Turkish and not null pronouns. For this reason, we have omitted the representation of the Turkish null pronoun in Figure 5.10.

pronoun should activate a topic interpretation more strongly than a non-topic interpretation, the Turkish overt pronoun should activate a non-topic interpretation more strongly than a topic interpretation. Hence, when listening to Dutch, the Turkish-Dutch bilingual child needs to inhibit the activation of the Turkish overt pronoun and its preferred non-topic interpretation. At the same time, she has to select the Dutch overt pronoun and its preferred topic interpretation.

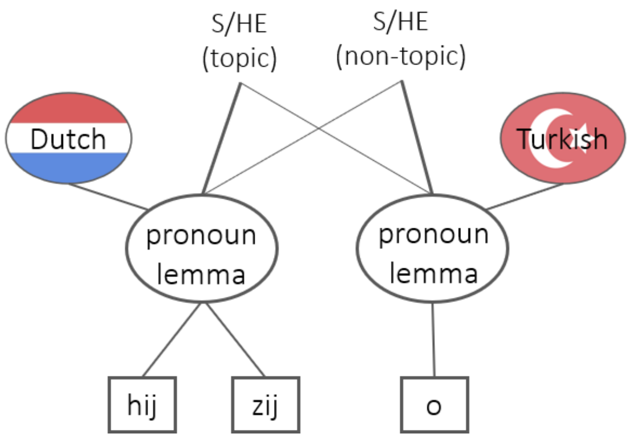


Figure 5.10. Schematic representation of lemmas for pronouns, their semantic representations and their forms in the language systems of Turkish-Dutch children. For simplicity, we only show the 3rd person singular forms of the Dutch pronouns *hij* and *zij* and the Turkish pronoun *o*. Following Kootstra & Doedens (2016), thicker lines represent stronger connections. The flags refer to the language membership of the lemmas (Dijkstra & van Heuven, 2002).[3]

The strength of language co-activation depends on language dominance, typically measured by bilinguals' proficiency in their two languages. The more dominant a bilingual is in the language not in use, the stronger this language becomes co-activated during the processing of the other language. In turn, more processing resources have to be allocated to inhibit this co-activation (Costa, Santesteban, & Ivanova, 2006; Hopp, 2017; Meuter & Allport, 1999; van Dijk, Dijkstra, et al., Chapter 3).

[3] We realize that our representation of pronoun lemmas and their semantic representations is a simplified one. Identifying the exact nature of pronoun lemmas is outside the scope of this thesis (see, e.g., Jurafsky, Bell, & Girand, 2002; Schmitt, Meyer, & Levelt, 1999, for a discussion of this topic).

The above account explains why Turkish-Dutch children fixated less on the Turkish-preferred disjoint referent with increased balance in their languages. The more balanced children were in their languages – as opposed to being Dutch-dominant – the more activation the Turkish overt pronoun structure received. Consequently, balanced Turkish-Dutch children had to suppress the co-activation of the Turkish overt pronoun and its preferred non-topic interpretation more strongly than Dutch-dominant children. Hence, Turkish-Dutch children were less likely to fixate on the disjoint referent the more balanced they were in their languages (also see Anderssen, Lundquist, & Westergaard, 2018, for a similar explanation of cross-linguistic influence in the speech production of adult heritage speakers). In this way, the above account explains why cross-linguistic influence from Turkish manifested itself in terms of fewer fixations on the Turkish preferred disjoint referent when the balance of bilingual children's languages increased.

The above account explains why Turkish-Dutch children fixated less on the Turkish-preferred disjoint referent with increased balance in their languages. The more balanced children were in their languages – as opposed to being Dutch-dominant – the more activation the Turkish overt pronoun structure received. Consequently, balanced Turkish-Dutch children had to suppress the co-activation of the Turkish overt pronoun and its preferred non-topic interpretation more strongly than Dutch-dominant children. Hence, Turkish-Dutch children were less likely to fixate on the disjoint referent the more balanced they were in their languages (also see Anderssen, Lundquist, & Westergaard, 2018, for a similar explanation of cross-linguistic influence in the speech production of adult heritage speakers). In this way, the above account explains why cross-linguistic influence from Turkish manifested itself in terms of fewer fixations on the Turkish preferred disjoint referent when the balance of bilingual children's languages increased.

There is some independent evidence for the account sketched here from an elicited production study on Ukrainian-English bilingual children's pronoun use in Ukrainian and English (Mykhaylyk & Ytterstad, 2017), and more specifically on how often children used a null *object* in their two languages. Null objects are allowed in Ukrainian in contexts where the object has already been mentioned in the discourse. In contrast, in English, null objects are ungrammatical. Hence, the situation for null and overt object use in their study resembles null and overt subject use in Turkish and Dutch. The bilingual children tested were aged between 4 and 6 years, living in the United States and Ukrainian was spoken at home by their parents. The authors found that 5- and 6-year-olds used more null objects in their Ukrainian than monolingual peers, even though this option is not allowed in English.

We believe that the behaviour of the Ukrainian-English bilingual children can be accounted for in terms of co-activation and inhibition. In our view, the intention of a Ukrainian-English bilingual child to produce an object pronoun structure in Ukrainian activates object pronoun representations in both Ukrainian and English. Hence, the child has to inhibit the co-activation of English and, in parallel, choose the preferred realisation of the object pronoun in Ukrainian (i.e., overt or null). Because the overt realisation of objects is strongly connected to the English object pronoun representation, the bilingual child might inhibit this strategy while inhibiting the activation of English pronouns. Consequently, overt object pronouns become less available for use in Ukrainian. This explains why children more often chose a null object structure. Furthermore, this view also accounts for why the 4-year-olds did not show an overproduction of null objects as compared to the 5- and 6-year-olds: on the assumption that they were more dominant in Ukrainian (because they had yet to start school), co-activation of the overt object structure in English should have been weaker and hence less inhibition would have been required. In this sense, they resemble the Dutch-dominant children in our study, and the 5 and 6 year olds were more similar to the balanced Turkish-Dutch children. This view can furthermore explain observations of cross-linguistic influence that go in the opposite direction than is typically predicted, so-called 'overcorrection' (Kupisch, 2014).

The mechanisms of co-activation and inhibition also offer an additional explanation for why effects of cross-linguistic influence were more pronounced online than offline. Inhibition effects of the Turkish overt pronoun structure and its preferred non-topic interpretation *online* result in fewer disjoint referent choices *offline*. This is because inhibited representations and interpretations are less likely to become selected, upon the assumption that the highest activated representation becomes selected (e.g., Levelt, Roelofs, & Meyer, 1999). Hence, children's offline choices should reflect inhibition online. At the same time, inhibition effects might be temporary and resolved before children make a choice offline. To be more precise, children may initially inhibit the non-topic interpretation when they encounter the overt pronoun and directly thereafter, only to later reconsider this interpretation. Hence, inhibition effects experienced *during* pronoun resolution might no longer affect children's offline choices. If this is correct, we would only expect the inhibition effects observed online to also surface offline when they are so strong that children are unable to resolve them. This would explain why more balanced children displayed more local referent fixations online and local referent choices offline, and why these effects were only strong enough online to reach significance.

One remaining question is how our account aligns with previous findings of cross-linguistic influence in offline pronoun resolution studies with bilingual children (e.g., Argyri & Sorace, 2007; Serratrice, 2007; Sorace et al., 2009) and in Schimke et al.'s (2018) online pronoun resolution study with bilingual adults, where effects of cross-linguistic influence went in the opposite direction: whilst children in our study inhibited pronoun preferences from Turkish in Dutch pronoun resolution, children and adults in these earlier studies were found to overaccept pronoun preferences from the one language into the other. For instance, Greek-English bilingual children in Argyri and Sorace (2007) accepted pragmatically infelicitous overt pronouns more often than their Greek monolingual peers, suggesting they used a strategy from English (i.e., always use an overt pronoun rather than a null pronoun) in their Greek. We believe that the explanation for the different outcomes between our and their studies should be sought in the dominance profiles of the participants studied.

Cross-linguistic influence in previous studies was typically obtained for participants who were dominant in the language not in use (e.g., Argyri & Sorace, 2007; Schimke et al., 2018; Sorace et al., 2009).[4] We therefore expect that in these participant groups the co-activation of the overt pronoun structure in the language not in use may have been so strong that children and adults were unable to fully inhibit these representations. This explains why children in Argyri and Sorace (2007) and Sorace et al. (2009) overused the overt pronoun structure from English into Greek and Italian and why adult L2 learners of German with Spanish as L1 in Schimke et al. (2018) fixated more on the non-topic referent in the German eye-tracking task than a bilingual control group. The children we tested were either dominant in Dutch or relatively balanced in their two languages. We would predict that Turkish-dominant children would show an increase in fixations on and in choices of the disjoint referent.

Future directions

Finally, there were some limitations to this study. First, participant groups were relatively small. We are currently collecting more data to test whether our observations hold with larger participant groups. Furthermore, we tested pronoun resolution for one type of sentence structure only, that is, *while-*

[4] Schimke and colleagues (2018) do not explicitly mention the dominance profiles of their participants. However, the participants' self-rated German proficiency (range = 2-4 with a maximal possible score of 5) suggests they were not at near-native level in their L2.

clauses. Given that the type of sentence structure can affect the presence or absence of (offline) effects of cross-linguistic influence in adult L2 learners (e.g., Contemori, Asiri, & Perea Irigoyen, 2019), future studies should replicate our design with different structures. Moreover, more studies are necessary for comparing online and offline effects of cross-linguistic influence in bilingual children. We believe that the account sketched in the above offers directions to researchers to formulate clear predictions about cross-linguistic influence in bilingual children's sentence processing, which hopefully aids future studies.

5.6 Conclusions

The present study was, to the best of our knowledge, the first to investigate cross-linguistic influence during online pronoun resolution in bilingual children. Using the visual world paradigm, we obtained evidence for cross-linguistic influence from Turkish into Dutch. In particular, the more dominant children were in their Turkish, as measured by their relative proficiency and exposure, the less they fixated on the non-topic antecedent when listening to Dutch pronouns. In contrast, no evidence was found for a general effect of bilingualism affecting children's online behaviour. Furthermore, whilst similar trends of cross-linguistic influence were observed online and offline, they only reached significance online.

We explained our results in terms of co-activation and inhibition following online studies with bilingual children and L2 adults (e.g., Hopp, 2017; van Dijk, Dijkstra & Unsworth, Chapter 3): the more Turkish-dominant children were, the stronger the non-topic antecedent became activated through Turkish, the greater the inhibition needed to suppress it. This account can also explain previously observed effects of cross-linguistic influence in offline and online pronoun resolution in bilingual children (e.g., Argyri & Sorace, 2007; Sorace et al., 2009) and L2 adults (e.g., Schimke et al., 2018). Moreover, the account correctly predicts stronger effects of cross-linguistic influence online than offline. Crucially, our study shows that online tasks are essential to better understand subtle effects of cross-linguistic influence in order to develop a comprehensive theory of cross-linguistic influence in bilinguals.

CHAPTER 6

Discussion

In this thesis, we investigated the occurrence of cross-linguistic influence at the level of morphosyntax during real-time sentence processing in simultaneous bilingual children and adults. An example of the topics we considered is whether the Turkish preference to assign overt pronouns a non-topic interpretation influenced Turkish-Dutch children's overt pronoun resolution in Dutch. Cross-linguistic influence is a well-attested phenomenon in bilingual children in their sentence production and offline comprehension, but research using online techniques is sparse. This thesis set out to answer five related questions:

- To what extent and in what manner does cross-linguistic influence manifest itself during sentence processing in bilingual children? (Chapters 3 and 5)
- To what extent is online cross-linguistic influence predicted by surface overlap between languages and by language dominance? (Chapters 3, 4 and 5)
- Is there evidence for general processing difficulties in bilingual children's sentence processing? (Chapters 3 and 5)
- How does cross-linguistic influence during sentence processing develop into adulthood? (Chapter 4)
- How do effects of cross-linguistic influence during real-time sentence processing relate to cross-linguistic influence in offline comprehension and production? (Chapters 2 and 5)

In the following section, we review the findings of production and offline comprehension studies on cross-linguistic influence in bilingual children (Chapter 2) to set the stage for a discussion of our online studies (Chapters 3, 4 and 5) in the subsequent section. In the second part of this chapter, we incorporate our findings within a model of sentence processing in bilingual children, explaining the mechanisms we deem responsible for cross-linguistic influence.

6.1 Effects of cross-linguistic influence in offline comprehension and production in the literature

In Chapter 2, we systematically reviewed available experimental studies on cross-linguistic influence in bilingual children's sentence production and comprehension. Furthermore, in a meta-analysis we analysed effect sizes of cross-linguistic influence across studies and tested effects of surface overlap, language dominance and age. We made four observations that motivate why further (online) studies are crucial to better understanding cross-linguistic influence.

Our first observation was that the weighted average effect size of cross-linguistic influence across studies was significantly larger than zero. This showed that cross-linguistic influence occurs in bilingual children's offline behaviour. However, the effect size of cross-linguistic influence varied largely across studies and most variation between studies was left unexplained. Moreover, we also observed behaviour in bilingual children that was inconsistent with cross-linguistic influence. To be more precise, in some studies bilingual children behaved in the opposite direction than predicted. For example, in a study by Mykhaylyk and Ytterstad (2017), Ukrainian-English bilingual children were found to omit *more* object pronouns in Ukrainian than monolingual peers, even though English does not allow null objects. As yet, there is no comprehensive theory that can account for the variation in effects of cross-linguistic influence.

Second, not only did studies differ in terms of their outcomes, but also in the way in which surface overlap was operationalized. Surface overlap refers to the amount of (morpho)syntactic overlap between languages (e.g., Hulk & Müller, 2000). We found that different studies used different definitions for surface overlap. Consequently, it was not very informative to directly compare effects of surface overlap between studies. Therefore, we defined and coded for three types of overlap situations ourselves (following e.g., Unsworth, 2003): *partial overlap*, *no overlap* and *complete overlap*. *Partial overlap* refers to a situation in which morphosyntactic structures in bilingual children's languages have the same form, but differ in the circumstances in which they can be used or how they should be interpreted. *No overlap* refers to a situation in which bilingual children's languages use different morphosyntactic structures. *Complete overlap* refers to a situation in which morphosyntactic structures in bilingual children's languages have the same form, use and meaning. We only compared effect sizes of cross-linguistic influence with *partial* and *no overlap*, as we reasoned that cross-linguistic influence cannot become visible in children's sentence production and offline comprehension with *complete overlap*. There were no significant differences between overlap situations. However, often studies provided insufficient information to define surface overlap from the bilingual child's point of view. As a result, it was impossible to draw any definite conclusions about the role of surface overlap.

We also considered the studies' operationalization of language dominance, which refers to the balance between bilingual children's languages. Different studies used different proxies for language dominance, including language proficiency, amount of language exposure, and societal status of children's languages. Furthermore, some studies combined different measures of language dominance to divide children into dominance categories. Due to this large variability in definition and operationalization, it was not feasible to directly compare dominance effects between studies. Therefore, we used societal language status (i.e., language spoken in the larger society or at home only) as an objective proxy of dominance in our analyses. This information was straightforwardly available in the studies in our dataset. On the one hand, the effect of cross-linguistic influence was present in bilingual children regardless of the societal status of their language. On the other hand, the effect of cross-linguistic influence was larger in the home language than in the societal language. Hence, we concluded that language dominance predicts the strength of cross-linguistic influence in children's sentence production and offline comprehension (e.g., Bosch & Unsworth, 2020; Foroodi-Nejad & Paradis, 2009). However, we could not tell from our dominance proxy which measure of language dominance (e.g., language proficiency, language exposure) best explained bilingual children's behaviour.

Third, it was unclear whether effect sizes attested in individual studies were truly reflections of cross-linguistic influence or whether they were – in part – driven by a more general bilingualism effect. In this context, *general bilingualism* refers to the notion that bilingual children's sentence processing could be negatively affected by them having to deal with two languages instead of one (e.g., Sorace, 2011; Sorace & Serratrice, 2009). Hence, in previous studies, effects of cross-linguistic influence might have been confounded with general bilingualism effects. It was impossible to test for general bilingualism effects, however, because most existing studies did not include a bilingual control group (but cf. Sorace, Serratrice, Filiaci, & Baldo, 2009).

Finally, our meta-regression on the effect of age on cross-linguistic influence showed that cross-linguistic influence persists in older children. Although there was an overall pattern for the effect of cross-linguistic influence to decrease with age, this pattern failed to reach significance. Hence, these findings suggest that cross-linguistic influence is not specific to young children only. Instead, cross-linguistic influence might be part and parcel of being bilingual. This is in line with – the few – studies that found cross-linguistic influence in early bilingual adults (e.g., Kupisch, 2012; Runnqvist, Gollan, Costa, & Ferreira, 2013).

In sum, Chapter 2 showed that whilst cross-linguistic influence significantly affects bilingual children's behaviour on morphosyntactic production and offline comprehension tasks, individual effects of cross-linguistic influence differ between and within studies. Some of the variation across effect sizes could be accounted for by the societal status of children's language tested. However, which direct measure(s) of language dominance was (or were) driving the effect of societal status remains unknown, because language dominance was operationalized differently from study to study. A similar issue was observed for surface overlap: studies defined surface overlap in different ways, making a direct comparison difficult. In addition, it is unclear to what extent effects of cross-linguistic influence in existing studies consist of a more general bilingualism effect.

To address the issues in the above, we used online experimental techniques to investigate effects of cross-linguistic influence during sentence processing in bilingual children (**Chapters 3** and **5**) and adults and adolescents (**Chapter 4**). In addition, we systematically assessed the role of surface overlap (**Chapters 3** and **4**) and of different measures of language dominance: relative language proficiency, relative current language exposure, and relative cumulative language exposure (**Chapters 3-5**). We chose online techniques, because these might provide a more direct measure of cross-linguistic influence in bilingual children than offline measures, given that recent theories have suggested that cross-linguistic influence arises during language processing (e.g., Nicoladis, 2006, 2012; Nicoladis, Rose, & Foursha-Stevenson, 2010; Serratrice, 2007, 2016; Sorace & Serratrice, 2009). To investigate this possible difference between online and offline data, we directly compared the outcomes between an online and offline task in **Chapter 5**. Furthermore, each experiment contained a bilingual control group (e.g., Serratrice, Sorace, Filiaci, & Baldo, 2009; Sorace et al., 2009) to test and control for a general effect of bilingualism (**Chapters 3-5**).

6.2 Online effects of cross-linguistic influence in this thesis

In both online experiments with bilingual children, we found evidence for cross-linguistic influence during sentence processing (Chapters 3 and 5). Furthermore, the observed direction of the online effect (i.e., inhibition) was similar across studies.

In the self-paced listening experiment (Chapter 3), we tested whether the presence of a Dutch word order either overlapping or not with English or German would affect English-Dutch and German-Dutch children's listening times, respectively. We did indeed obtain evidence for cross-linguistic influence for overlapping word orders in the German-Dutch children. To be more precise, when German-Dutch children listened to Dutch long passive and Verb Second (V2) structures that overlapped partially or completely with German, their listening pace slowed down compared to the English-Dutch and monolingual children. Hence, we found evidence for inhibition effects during listening.

In the eye-tracking task (Chapter 5), we investigated whether Dutch overt pronoun processing was affected by Turkish in Turkish-Dutch bilingual children. In Dutch, overt pronouns typically refer back to the topic of the discourse (e.g., Ariel, 2014; Cardinaletti & Starke, 1999), but in Turkish they typically refer to a non-topic (e.g., Azar, Özyürek, & Backus, 2020; Enç, 1986). Children were less likely to consider the Turkish-preferred non-topic antecedent when listening to overt pronouns. Hence, in line with the self-paced listening task, the eye-tracking data suggest that cross-linguistic influence resulted in inhibition effects during sentence processing.

Our studies were among the first to investigate cross-linguistic influence in bilingual children using online techniques (cf. Lemmerth & Hopp, 2019) and, to the best of our knowledge, the first to observe cross-linguistic influence during real-time sentence processing in this population. The direction of the results suggests that bilingual children were keeping their languages apart during sentence processing by inhibiting options that were available in both the language being processed and the language not in use. Importantly, our findings indicate that cross-linguistic influence is not limited to children's speech production and offline comprehension (e.g., Serratrice, 2013; Chapter 2), but affects sentence processing as well. Furthermore, the findings are in line with those found for adult L2 learners (e.g., Foucart & Frenck-Mestre, 2012; Hopp, 2017) as well as with observations of non-convergent effects of cross-linguistic influence in bilingual children in Chapter 2 and overcorrection in studies on simultaneous bilingual adults' sentence production and grammaticality judgements (e.g., Anderssen, Lundquist, & Westergaard, 2018; Kupisch, 2014).

Predictors of online cross-linguistic influence

In our online studies, surface overlap and language dominance moderated effects of cross-linguistic influence during sentence processing in bilingual children. In addition, we identified a third, unexpected predictor of online cross-linguistic influence: lexical overlap. We will now discuss the observed effects of these three predictors one by one.

Surface overlap

In our self-paced listening task (Chapter 3), we systematically assessed the role of surface overlap in bilingual children by comparing effects of cross-linguistic influence in situations of *partial overlap*, *no overlap* and *complete overlap*. We found cross-linguistic influence to be most pronounced in situations of *partial overlap*. In particular, German-Dutch children slowed down when listening to partially overlapping long passive sentences in Dutch compared to the other two groups. Cross-linguistic influence arose as well in a situation of *complete overlap* (V2), but only in relation to language dominance (see next section). In contrast, cross-linguistic influence was not attested when there was *no overlap* between bilingual children's languages.

Our findings are in line with existing production and offline comprehension studies involving bilingual children (e.g., Hulk & Müller, 2000; Müller & Hulk, 2001) and online studies involving adult L2 learners (e.g., Foucart & Frenck-Mestre, 2012; Hopp, 2017) that observed cross-linguistic influence with *partial overlap* but not with *no overlap*. Our findings are also in line with online studies with adult L2 learners that observed cross-linguistic influence with *complete overlap* (e.g., Alemán Bañón et al., 2014). At the same time, our findings differ from studies with children that observed cross-linguistic without surface overlap (e.g., Argyri & Sorace, 2007; Foroodi-Nejad & Paradis, 2009) and our meta-analysis in which *partial overlap* had no special status (Chapter 2). We believe this discrepancy between results should be sought in our use of an *online comprehension* task versus *production* and *offline comprehension* tasks typically employed in previous studies. To be more precise, cross-linguistic influence during sentence processing might be less relevant when bilingual children's languages do not share the structure being processed. We return to this point in Section 3 below.

Language dominance

We systematically assessed effects of language dominance on online cross-linguistic influence in bilingual children in our self-paced listening experiment (Chapter 3) and our eye-tracking experiment (Chapter 5). Language dominance was operationalised using three measures: relative language proficiency, relative current language input, and relative cumulative language input. In both experiments, we found language dominance to affect the strength of cross-linguistic influence. In particular, the more proficient children were in the language not in use (i.e., German and Turkish) compared to the language being tested (Dutch), the more they slowed down when processing overlapping word orders in the self-paced listening task and the less they fixated on the non-topic antecedent in the eye-tracking task. Crucially, some effects of cross-linguistic influence only emerged when language dominance was taken into account. This was the case for the completely overlapping word order between Dutch and German in the self-paced listening task (V2) and for pronoun resolution in the eye-tracking task. Hence, our findings show that it is essential to take into account language dominance in order to be able to detect subtle effects of cross-linguistic influence online.

In all, our findings are in line with previous studies that found language dominance to affect the presence and the strength of cross-linguistic influence in children's sentence production and offline comprehension (e.g., Bosch & Unsworth, 2020; Foroodi-Nejad & Paradis, 2009; Kidd, Chan, & Chiu, 2015). They also agree with online studies on bilingual adults that found cross-linguistic influence to decrease or disappear the more dominant participants were in the language tested, as measured by proficiency (e.g., Foucart & Frenck-Mestre, 2012; Hopp, 2017; Martohardjono, Phillips, Madsen II, & Schwartz, 2017).

Interestingly, our results suggest that relative proficiency might better predict online cross-linguistic influence than exposure measures do. To be more precise, whilst similar trends were observed in our experiments for children's relative current and cumulative language exposure as for relative proficiency, the input measures often failed to reach significance and were less stable over the different conditions. This difference between measures cannot be straightforwardly explained by, for example, differences in variance: ranges are comparable for our relative proficiency and input measures. In section 3, we will explore why relative proficiency might be a better predictor of cross-linguistic influence than relative language exposure. Importantly, our observations support the conclusion from our meta-analysis that it is important to use different proxies of language dominance (also see Unsworth, Chondrogianni, & Skarabela, 2018).

Lexical overlap

Finally, in our self-paced listening experiment (Chapter 3) we identified a third possible predictor of online cross-linguistic influence, namely, lexical overlap. We used this variable to explain why we observed cross-linguistic influence in the German-Dutch children, but not in the English-Dutch children. We argued that because German is more closely related to Dutch than English (e.g., Schepens, van der Slik, & van Hout, 2013), cross-linguistic influence during sentence processing was more likely to be observed in German-Dutch bilinguals than English-Dutch bilinguals. Indeed, more than two thirds of the Dutch words at the critical regions of the sentence tested were cognates with German and only about half of the words were cognates with English. A relation between lexical overlap and cross-linguistic influence during sentence reading – although in the opposite direction – has been observed in L2 learners (Hopp, 2017). We are not aware of any study on cross-linguistic influence with bilingual children that has assessed the role of lexical overlap.

General bilingualism effects on online processing

It has been argued that bilingualism can result in less efficient sentence processing (e.g., Sorace, 2011). However, we found no evidence of such general processing difficulties in children in either of our online experiments (Chapters 3 and 5). In fact, in the self-paced listening task, the bilingual groups processed sentences significantly faster than their monolingual peers. Furthermore, in the eye-tracking study, Turkish-Dutch and German-Dutch bilingual children showed overall similar fixation patterns as their monolingual peers. Where the two groups of bilingual children were observed to behave differently from monolingual children, they were found to behave differently from each other as well. In other words, we observed no negative effects due to bilingualism during processing. Instead, the observation of faster processing behaviour in bilingual children in the self-paced listening experiment suggests that *if* being bilingual affects sentence processing in children, it results in more efficient sentence processing. More research is needed to explore this observation further.

Our findings contrast with the results of online studies with adult L2 learners, which did observe negative general bilingualism effects (e.g., Clahsen & Felser, 2006; Cunnings, Fotiadou, & Tsimpli, 2017; Hopp, 2010; Roberts, Gullberg, & Indefrey, 2008), suggesting that processing difficulties observed in adult L2 learners should not be automatically extended to other bilingual populations (also see Felser, 2020). Note, however, that the children in both our experiments were tested in their societal language. As a consequence, most children were either dominant in the language tested or relatively balanced in both of their languages. It therefore remains unclear whether the absence of evidence for general bilingualism effects during sentence processing extends to children who are dominant in their other language. Future studies should investigate processing behaviour in children's non-dominant language.

Comparing findings of cross-linguistic influence online and offline

In Chapter 5, we directly compared children's offline effects of cross-linguistic influence in a picture selection task to effects observed during pronoun resolution in online comprehension. We found similar patterns in children's online eye-tracking and offline comprehension data: The more dominant children were in Turkish (as measured by their relative proficiency in Dutch and Turkish), the less likely they were to choose a non-topic antecedent for overt pronouns. This finding reflected children's online fixations on the non-topic antecedent. At the same time, however, offline effects were less pronounced than online effects. In particular, while the observed online effects for language dominance were significant, the offline effect of language dominance failed to reach significance.

The absence of a significant effect of offline cross-linguistic influence in our data follows previous studies that observed numerical but non-significant trends of cross-linguistic influence in children's sentence production and offline comprehension (e.g., Argyri & Sorace, 2007; Nicoladis, 2003; Serratrice, Sorace, Filiaci, & Baldo, 2012). As discussed in Chapter 2, studies with bilingual children are often underpowered. This might explain the absence of significant results. The presence of clear and significant effects in our eye-tracking study with the same (limited) number of participants as in the offline task, suggests that online studies might be better suited to study subtle effects of cross-linguistic influence. A similar conclusion was drawn in a study with early bilingual adults (e.g., Martohardjono et al., 2017). Hence, we found support for the view that online techniques are essential tools to investigate cross-linguistic influence in bilingual children, in addition to offline techniques.

Online cross-linguistic influence in adolescence and adulthood

Finally, in Chapter 4, we investigated online effects of cross-linguistic influence in simultaneous bilingual adults and adolescents. In this study, we replicated the self-paced listening experiment in Chapter 3 with an older generation of bilingual and monolingual children as participants. We observed evidence for online cross-linguistic influence in German-Dutch bilingual adults and adolescents. Similar to the child study, cross-linguistic influence was manifested as a slowdown effect in overlapping word orders. In contrast to the child study, however, cross-linguistic influence was attested for adults and adolescent only in a situation of *partial overlap* in interaction with language dominance, and not in a situation of *complete overlap*. Furthermore, the effect was only observed when the experiment was in a bilingual language mode and not in a monolingual language mode, as was the case for the children.

Our findings suggest that cross-linguistic influence persists into adulthood and that its effects are qualitatively similar in children and adults. At the same time, however, our results show that online cross-linguistic influence is quantitatively different in the two groups, that is, less pronounced in adults than in children. Our findings are in line with a number of online studies supporting cross-linguistic influence in early bilingual adults as well (e.g., Kupisch, 2012, 2014; Martohardjono et al., 2017; Runnqvist et al., 2013). Our study adds to the existing body of knowledge by directly comparing online results in bilingual adults to bilingual children.

To sum up, in this thesis we found evidence for cross-linguistic influence during sentence processing in bilingual children. Such influence was manifested as an inhibition effect. Furthermore, we observed that the degree of surface overlap, lexical overlap, and language dominance moderated the effect of cross-linguistic influence. We found no evidence that bilingualism in general results in less efficient sentence processing. On the contrary, our self-paced listening data suggested that bilingual children were overall more efficient sentence processers than their monolingual peers. In addition, our direct comparison of children's offline and online sentence comprehension suggests that online techniques are more suitable to study effects of cross-linguistic influence in bilingual children than offline techniques. Finally, our replication of the self-paced listening study suggests that qualitatively similar effects of cross-linguistic influence during sentence processing are at play in simultaneous bilingual children, adults and adolescents, although there might be quantitative differences between the groups. Taken together, our findings show that online studies are essential to understand cross-linguistic influence in bilingual acquisition.

6.3 A model of sentence processing in bilingual children and adults

On the basis of our findings, we are able to formulate a detailed theoretical model that can explain the online cross-linguistic influence observed in bilingual children and adults: the Cross-Linguistic Influence during Sentence Processing (**CLISP**) model. This model is based on existing models of sentence processing and specifically on models of cross-linguistic influence and priming in bilingual children and adults discussed in the introduction (e.g., Hartsuiker, Pickering, & Veltkamp, 2004; Hartsuiker & Bernolet, 2017; Nicoladis, 2006, 2012; Nicoladis et al., 2010; Serratrice, 2016). Because these models focus on language production rather than on comprehension, they cannot straightforwardly account for the inhibition effects reported in this thesis, although they do predict the presence of structural co-activation between bilingual's languages. Therefore, to accommodate our findings, we extend the models to account for sentence processing in bilingual children and adults as well. Central to the extended model are the concepts of language co-activation, inhibitory control and working memory resources (e.g., Conway & Engle, 1994; Costa & Santesteban, 2004; Green, 1998; Hopp, 2017; Meuter & Allport, 1999).

We first describe the relevant components of CLISP, before illustrating how it can account for the cross-linguistic influence we observed during the processing of (i) Dutch long passives with a post-verbal by-phrase in German-Dutch children (Chapter 3; PP-V structures); and (ii) Dutch pronouns in Turkish-Dutch children (Chapter 5). Next, we will illustrate how effects of surface overlap (Chapter 3) and language dominance (Chapters 3 and 5) follow from the model. Furthermore, we show how the model can also account for cross-linguistic influence in children's sentence production and offline comprehension (Chapters 2 and 5). Finally, we extend CLISP to account for the presence and absence of online cross-linguistic influence during sentence processing in simultaneous bilingual adults (Chapter 4).

Model components

We first discuss four components of the CLISP model we deem crucial to account for online effects of cross-linguistic influence in bilingual children: (a) representations and their interconnections; (b) processing; (c) cognitive control; and (d) memory resources. Components (a) and (b) draw heavily on existing sentence processing models and in particular the bilingual structural priming model by Hartsuiker and colleagues based on Pickering and Branigan (1998; e.g., Hartsuiker et al., 2004; Hartsuiker & Bernolet, 2017). Components (c) and (d) are added and represent mechanisms we deem essential to account for cross-linguistic influence in bilingual sentence processing.

(a) Representations and their interconnections

In constructing our model, we assume that during sentence processing at least three structural levels of representations (or *frames,* see Dell, 1986) are involved: (i) a phonological level, associated with spoken word forms; (ii) a morphosyntactic level, associated with syntactic structures and word-like representations called lemmas (Kempen & Huijbers, 1983)[1]; and (iii) a conceptual level, associated with word meanings and their interrelations (e.g., Ferreira & Dell, 2000; Levelt, 1989; Levelt, Roelofs, & Meyer, 1999). At the phonological level, speech or written text is converted into meaningful (sub)lexical units, such as syllables and words. At the morphosyntactic level, morphosyntactic structures are created with slots for lemmas, which are activated by lexical phonological units. At the conceptual level, information from the morphosyntactic level is mapped onto an event structure, representing word meaning and its interrelations. In sentence production models, such as models on cross-linguistic influence and cross-language priming, morphosyntactic properties of words are stored at or around the lemma level (e.g., Levelt et al., 1999). Lemmas are linked to combinatorial nodes that specify the grammatical structures in which a lemma can be used (e.g., Pickering & Branigan, 1998; Schoonbaert, Hartsuiker, & Pickering, 2007, but cf. Vosse & Kempen, 2009). Furthermore, lemmas are connected to their semantic representations (e.g., Levelt et al., 1999; Roelofs, 1992). In line with recent models of bilingualism, we assume that the lemma level plays a central role in cross-linguistic influence in bilingual children (e.g., Hartsuiker et al., 2004; Hartsuiker & Bernolet, 2017; Nicoladis, 2006, 2012; Nicoladis et al., 2010).

A schematic example of representations and connections at the lemma level is given for the Dutch passive auxiliary *worden* ("to be") in Figure 6.1. For the sake of simplicity, we assume that *worden* only occurs in combination with a main verb and a prepositional phrase in the two word orders tested in our self-paced listening studies (Chapters 3 and 4). An example of a PP-V and V-PP structure is given in (1) and (2), respectively. The lemma for *worden* is thus linked to two combinatorial nodes: (aux)V-PP and (aux)PP-V. The connection with the V-PP combinatorial node is stronger than the connection with the PP-V node, under the assumption that the V-PP structure is more frequent than the PP-V structure (e.g., Bernolet, Hartsuiker, & Pickering, 2009). This aspect is indicated in the figure by a thicker line (following e.g., Kootstra & Doedens, 2016). Furthermore, the lemma is also linked to its semantic representation.

[1] Following the language production literature, we will call the word-like units associated with morphosyntactic structures 'lemmas', although the involvement of lemmas is not often made explicit in the comprehension domain.

(1) **PP-V**
De leeuw wordt **door de beer geduwd**.
the lion is being by the bear pushed
"The lion is being pushed by the bear."

(2) **V-PP**
De leeuw wordt **geduwd door de beer**.
the lion is being pushed by the bear
"The lion is being pushed by the bear."

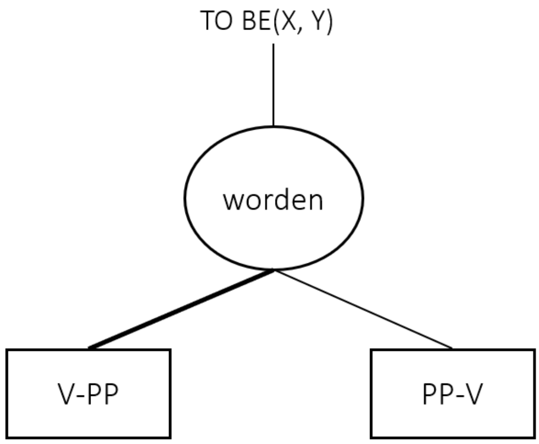


Figure 6.1. Schematic representation of the lemma for *worden* ("to be"). The lemma *worden* is linked to a semantic representation (TO BE) and to combinatorial nodes (V-PP and PP-V).

Evidence suggests that syntactic representations that are similar in form between languages are shared when a participant's proficiency is sufficient (e.g., Hartsuiker & Bernolet, 2017; Hartsuiker et al., 2004). This means, for instance, that German-Dutch children have a lemma for the German passive auxiliary *werden* in addition to the Dutch lemma *worden*. Both lemmas are indirectly connected through the semantic representation and the combinatorial nodes that they share (see Figure 6.2, adapted from Hartsuiker et al., 2004). Furthermore, we follow the assumption from Hartsuiker and colleagues (2004) that only the lemmas and not their semantic representations and combinatorial nodes are tagged for language.

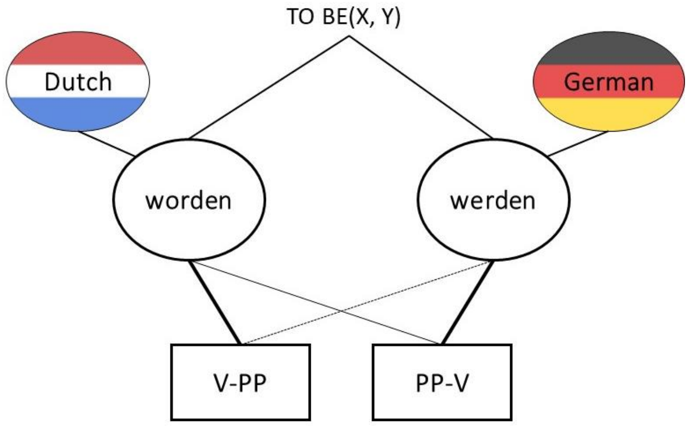


Figure 6.2. Schematic representation of the Dutch passive auxiliary *worden* and the German passive auxiliary *werden* at the lemma level. Note: The connection between *werden* and the V-PP node is represented by a dotted line. This is because the V-PP structure is officially ruled out in German grammar, but is to a certain extent present in spoken language (e.g., Betz, 2008; Dürscheid, 2012; Haider, 2010). This suggests that, for some speakers at least, German has a structural representation for the V-PP structure.

(b) Processing

During sentence processing, activation flows from lemmas to connected semantic and combinatorial nodes (e.g., Dell, 1986; Hartsuiker et al., 2004; Levelt, 1989; Levelt et al., 1999; Pickering & Branigan, 1998). Hence, in our Dutch example, when encountering the passive auxiliary *wordt* ("is being") in (1) and (2), the listener activates the lemma for *worden*. In turn, the V-PP and PP-V nodes and the semantic representation of *worden* become activated. As a result, the listener might predict a V-PP or PP-V structure coming up. Importantly, activation has been argued to flow in the opposite direction as well: from combinatorial nodes and semantic representations to lemmas (e.g., Dell, 1986). Consequently, the activation of the Dutch lemma *worden* results in the co-activation of the German lemma *werden* through their syntactic and semantic nodes.[2] Indeed, there is ample evidence of such lexical and syntactic co-activation in bilingual adults and – albeit to a lesser extent – in bilingual

[2] Note that *worden* and *werden* can also directly activate each other due to their overlap in form (e.g., Dijkstra et al., 2019). We would thus predict increased language co-activation with cognates (also see Chapter 3). Because the role of cognate status during sentence processing was not a focus of this thesis, we will not consider this matter any further.

children (e.g., Dijkstra & van Heuven, 2002; Dijkstra et al., 2019; Loebell & Bock, 2003; Vasilyeva et al., 2010; Von Holzen & Mani, 2012). As a result, lexical representations from different languages compete for selection during sentence processing.

(c) Cognitive control

In order to manage (co-)activation and competition between languages and to select the appropriate representations, the listener has to exert cognitive control (e.g., Green & Abutalebi, 2013). Control involves at least four mechanisms: decision, inhibition, shifting/switching, and updating (e.g., Green, 1998; Miyake et al., 2000). In our discussion we focus on the role of inhibition, following studies on lexical processing in bilinguals (e.g., Costa & Santesteban, 2004; Green, 1998), although we do not rule out the involvement of the other mechanisms as part of an explanation of online cross-linguistic influence.

With respect to inhibition, it has been proposed that inhibitory control processes are necessary to suppress spreading co-activation to non-target language representations (e.g., Green, 1998; Green & Abutalebi, 2013). Inhibition can be proactive and, especially relevant to the present model, reactive: proactive by suppressing the activation of a language in general before actually producing speech (e.g., Wu & Thierry, 2017), and reactive by suppressing the co-activation of non-target representations (e.g., Green, 1998). This implies that inhibitory control processes are necessary, for instance, to suppress the co-activation of the German lemma *werden* while processing the Dutch long passive structure.

The strength of inhibition observed during bilingual sentence processing depends on bilinguals' proficiency and daily exposure to their languages. In particular, the more proficient bilinguals are in one of their languages, the more strongly this language needs to be inhibited while processing their other language (e.g., Costa & Santesteban, 2004; Meuter & Allport, 1999). Furthermore, recent evidence suggests that a similar effect holds for language exposure: The more bilinguals are exposed to a language on a daily basis, the stronger it needs to be inhibited while processing the other language (e.g., Bonfieni, Branigan, Pickering, & Sorace, 2019).

(d) Working memory resources

Finally, sentence processing relies on working memory resources. Following Just and Carpenter (1992), we assume that working memory as it relates to language comprehension consists of three components: (i) the temporal storage of incoming sentence information; (ii) the temporal storage of sufficiently activated items from the long-term memory, such as lemmas; and (iii) other language comprehension processes, such as those predicting upcoming sentence information. Each component takes up working memory resources during sentence processing. Such resources are limited and when the capacity of working memory is reached, less strongly activated information will be de-activated and processes will be performed more slowly (e.g., Just & Carpenter, 1992).

In short, the more non-target representations become activated during sentence processing, the more working memory resources are necessary to temporally store these representations. Inhibitory control (and other cognitive control mechanisms) may, therefore, be essential to manage the amount of activation in working memory. However, at the same time, inhibitory processes take up memory resources as well (e.g., Conway & Engle, 1994; Just & Carpenter, 1992). Hence, it follows that if a listener's maximum working memory capacity is reached during sentence processing with multiple non-target representations still co-activated in working memory, sentence processing will slow down and an incorrect representation might become selected (e.g., Hopp, 2017).

Slowdown effects and fixations

We now turn to observations of cross-linguistic influence in this thesis in relation to the CLISP model. We will focus on the observed online cross-linguistic influence in the German-Dutch bilingual children's processing of long passives with a PP-V structure and the Turkish-Dutch bilingual children's online pronoun resolution. In the following sections, we explain how these findings can be accounted for by CLISP.

Long passives in German-Dutch bilingual children

An example of a PP-V structure German-Dutch bilingual children listened to in the self-paced listening task (Chapter 3) is given in (1), repeated below. Recall that we observed that the German-Dutch bilingual children slowed down compared to Dutch monolingual children directly after having heard the main verb (*geduwd*, "pushed"). This, we believe, is straightforwardly explained by CLISP.

(1) **PP-V**
De leeuw wordt door de beer geduwd.
the lion is being by the bear pushed
"The lion is being pushed by the bear."

The processing of the PP-V structure over time is schematically shown in Figure 6.3.A and 6.3.B below. First, the child hears *de leeuw* ("the lion"). This activates the Dutch lemma for *leeuw* and, in turn, the German lemma for *Löwe* (see Figure 6.3.A) through their shared semantic representation. We assume, for simplicity, that the S-V structure becomes activated with *de leeuw* in subject position. Second, the child hears *wordt* ("is being"). This activates the Dutch lemma *worden*. Furthermore, the PP-V and V-PP nodes become activated. As a result, the German lemma *werden* receives activation as well through the shared semantic and syntactic nodes with *worden*. The child places *wordt* in the auxiliary position and assigns the thematic patient role to *de leeuw*. Third, the child hears *door de beer* ("by the bear"). This activates the Dutch and German lemmas for *bear* (*beer* and *Bär*). In addition, the PP-V structure is strongly activated, because the PP appears before the main verb. Hence, the child can select the PP-V structure and assign the thematic patient role to *de beer*. Finally, the child hears the main verb *geduwd* ("pushed"). This activates the Dutch *duwen* and indirectly the German lemma *schieben* ("to push"). Furthermore, the order of the verb confirms the PP-V order of the sentence. The child can enter the main verb in the sentence structure and the sentence can now be interpreted as an event during which the lion is being pushed by the bear.

Why did the German-Dutch children slow down when processing the PP-V structure? There are two explanations that are non-necessarily mutually exclusive. On the first explanation, the PP-V structure was less accessible to German-Dutch children because they inhibited co-activation from German. More specifically, upon encountering the passive auxiliary *wordt*, the German-Dutch children had to suppress the co-activation of the German lemma for *werden*. Consequently, activation of the combinatorial node connected to *werden* was also suppressed. Hence, the activation of the PP-V node decreased, making it more difficult to select. As a result, children showed a processing delay.

On the second explanation, the activation and selection of the PP-V node increased the activation of the lemmas it is connected to. Importantly, this holds for both Dutch and German. As a consequence, German-Dutch children had to allocate additional working memory resources to inhibit the co-activation of German lemmas (compared to a situation without syntactic overlap between languages). This resulted in a temporal shortage of working memory resources, and, in turn, in a processing delay (e.g., Just & Carpenter, 1992). In support of this second explanation, children in our experiment had longer listening times, the smaller their working memory span was as measured by a digit span task (e.g., King & Just, 1991).

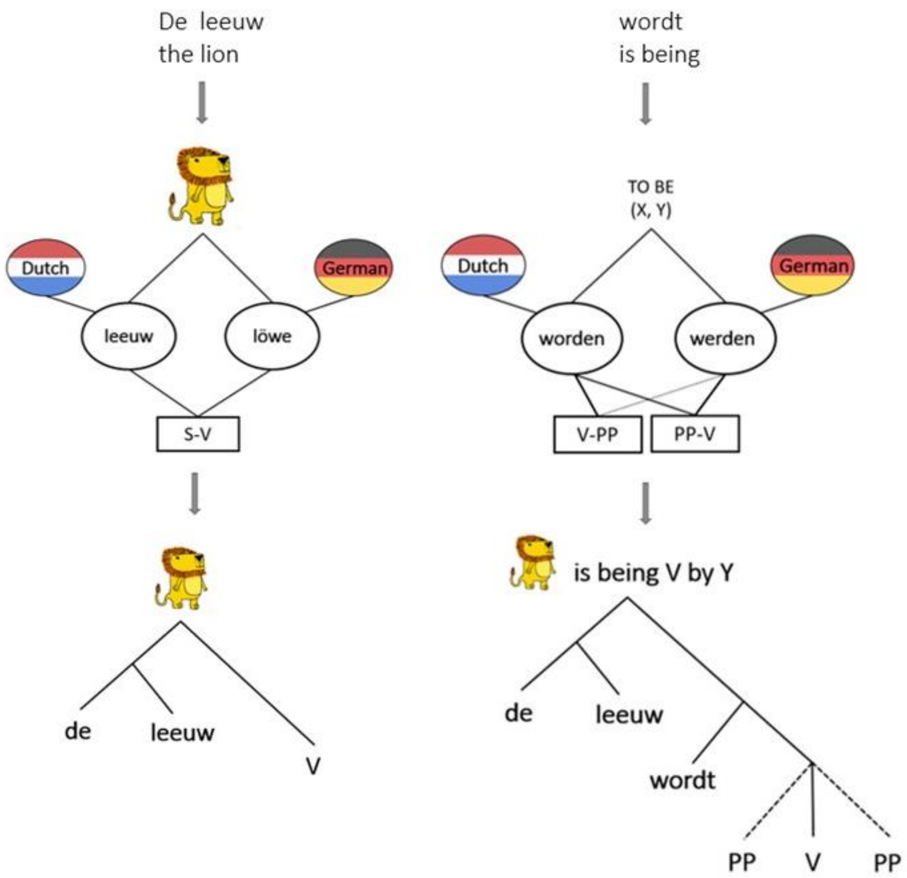


Figure 6.3.A. Schematic representation of the processing of a Dutch PP-V structure by a German-Dutch child. The figure shows the different constituents of the structure *de leeuw wordt door de beer geduwd* ("the lion is pushed by the bear"); the lemma representation for the lexical items with their combinatorial nodes; and the syntactic structures the child constructs from left to right with the thematic roles for the noun phrases.

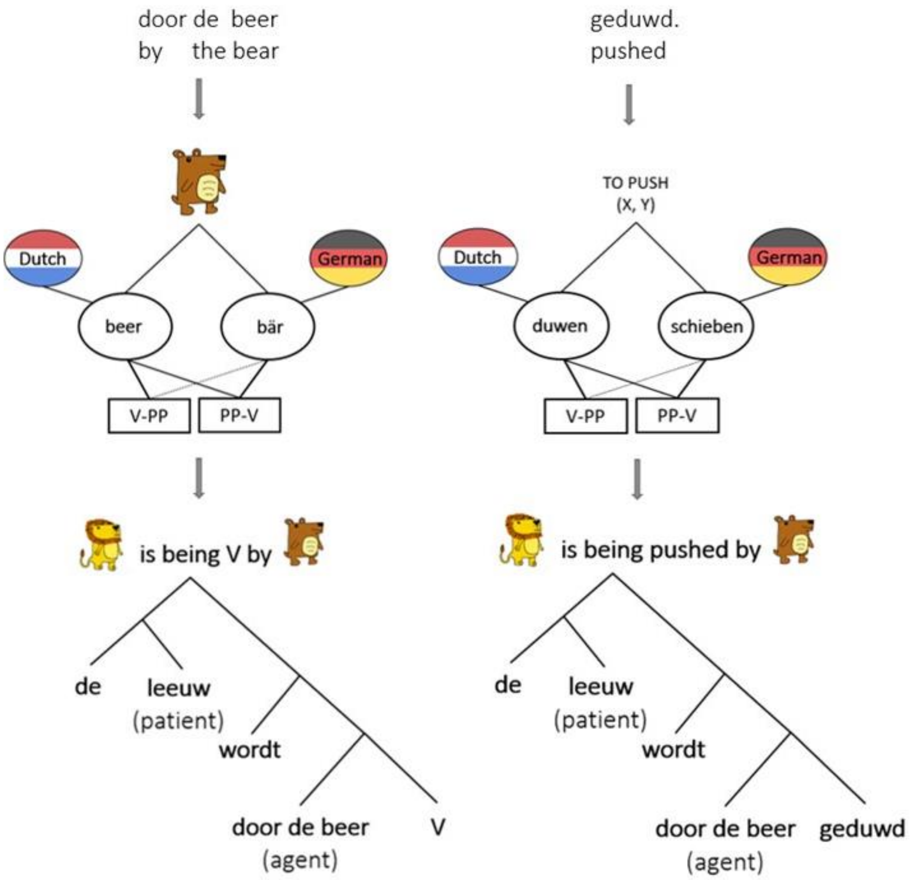


Figure 6.3.B. Schematic representation of the processing of a Dutch PP-V structure by a German-Dutch child. The figure shows the different constituents of the structure *de leeuw wordt door de beer geduwd* ("the lion is pushed by the bear"); the lemma representation for the lexical items with their combinatorial nodes; and the syntactic structures the child constructs from left to right with the thematic roles for the noun phrases.

Pronoun resolution in Turkish-Dutch bilingual children
An example of an overt pronoun structure tested in our eye-tracking study in Chapter 5 is given in (3). Recall that we found that Turkish-Dutch bilingual children were less likely to fixate on the non-topic referent, Sophie, the more balanced they were in their two languages. Again, this observation follows from the sentence processing model outlined above.

(3) Anna_i en Sophie_k zijn in de bibliotheek.
Anna_i and Sophie_k are in the library
Terwijl Anna_i een boek leest, neemt $\text{zij}_{i/k}$ een slokje water.
while Anna_i a book reads drinks $\text{she}_{i/k}$ a sip water

"Anna_i and Sophie_k are in the library. While Anna_i is reading a book, $\text{she}_{i/k}$ is taking a sip of water."

The processing of the overt pronoun sentence in (3) is schematically shown in Figure 6.4.[3] First, the child hears the verb *neemt* ("takes"). This activates the Dutch lemma *nemen* ("to take"), its meaning, and the syntactic representations in which it can occur. Furthermore, the Turkish translation equivalent of *nemen*, *alıyor* becomes activated. Because *nemen* appears in sentence-initial position the verb is placed in a VSO structure. Second, the child hears the overt pronoun *zij* ("she"). This activates the Dutch pronoun lemma and its semantic representations: referring back to a topic or a non-topic in the sentence. In addition, the Turkish overt pronoun *o* receives activation as well, through the semantic representations.[4] The overt pronoun *zij* is inserted in subject position in the VSO structure. Finally, the child hears *een slokje water* ("a sip of water"). This activates the Dutch and Turkish lemmas for *sip* and *water* – for simplicity, we only depicted the lemma for water in Figure 6.4. In turn, *een slokje water* is placed in object position.

[3] We realize that our representation of pronoun lemmas and their semantic representations is a simplified one. Identifying the exact nature of pronoun lemmas is outside the scope of this thesis (see, e.g., Jurafsky, Bell, & Girand, 2002; Schmitt, Meyer, & Levelt, 1999, for a discussion of this topic).
[4] Because the processing of overt pronouns activates discourse-pragmatic principles relevant to overt pronoun resolution (e.g., Serratrice, 2007), we assume that the processing of overt pronouns in Dutch activates only overt pronouns and their interpretations in Turkish and not null pronouns. We therefore do not consider null pronouns and their semantic representations in Turkish.

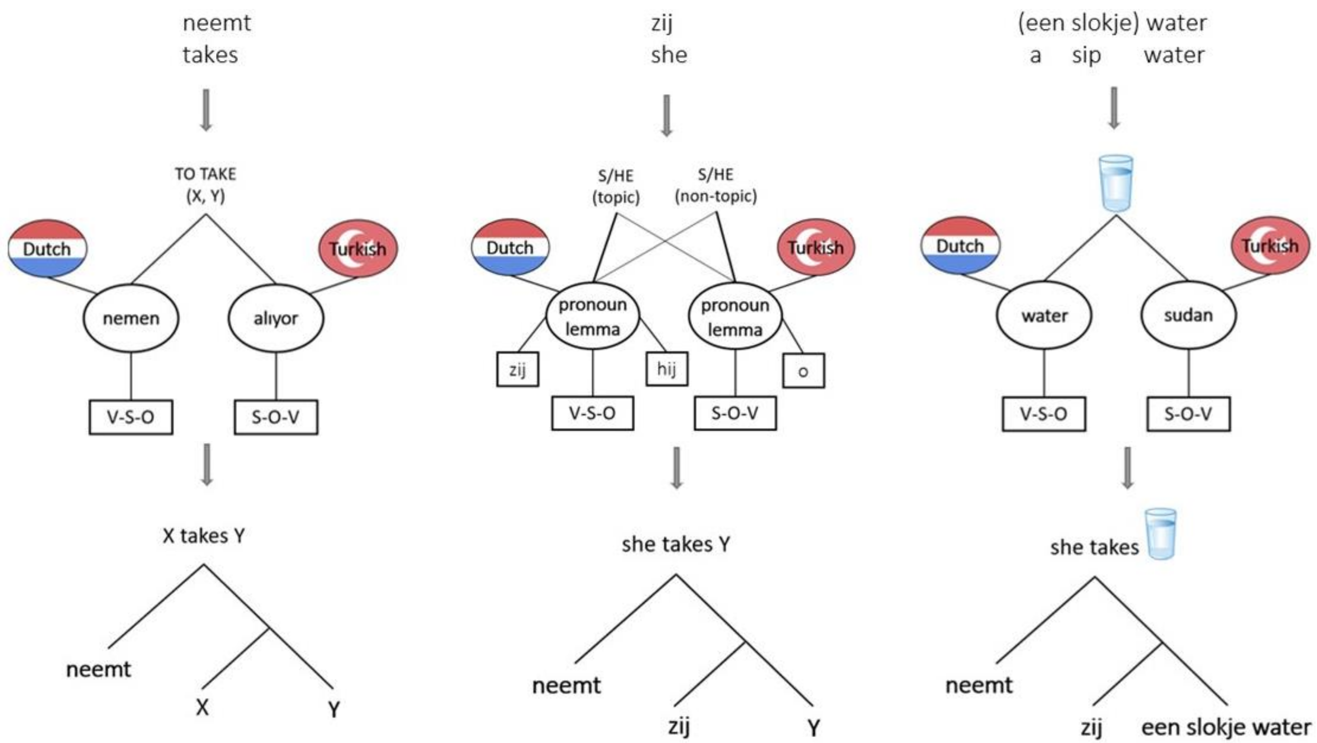


Figure 6.4. Schematic representation of the processing of a Dutch overt pronoun structure by a Turkish-Dutch child. The figure shows the different constituents of the clause *neemt zij een slokje water* ("she takes a sip of water"); the lemma representation for the lexical items with their combinatorial nodes and the pronoun forms *zij*, *hij* and *o*; and the syntactic structures the child constructs from left to right.

Again, the sentence processing model can straightforwardly explain why children's knowledge of Turkish resulted in fewer looks to Sophie, the non-topic referent, and more looks to Anna, the topic referent. While processing the Dutch sentences, Turkish-Dutch children had to suppress the co-activation of the Turkish overt pronoun *o*. In turn, this resulted in the inhibition of its associated meanings. Because the non-topic interpretation of an overt pronoun is preferred in Turkish over a topic interpretation, the activation of the non-topic interpretation was most strongly suppressed. In other words, the interpretation of *zij* as Sophie became temporally less available during Dutch sentence processing due to the inhibition of Turkish. The stronger such inhibition effects, the more Turkish-Dutch children fixated on Anna instead. Moreover, the non-topic interpretation was also more difficult to access later on in the sentence. This was evidenced by a decrease in looks to Sophie after the overt pronoun in sentences in which the gender on the pronoun forced a non-topic interpretation, such as in (4).

(4) Joris$_i$ en Sophie$_k$ zijn in de bibliotheek.
Joris$_i$ and Sophie$_k$ are in the library
Terwijl Joris$_i$ een boek leest, neemt zij$_k$ een slokje water.
while Joris$_i$ a book reads drinks she$_k$ a sip water

"Joris$_i$ and Sophie$_k$ are in the library. While Joris$_i$ is reading a book, she$_k$ is taking a sip of water."

In sum, the interaction between language co-activation, inhibitory control mechanisms, and working memory resources accounts for slowdowns in German-Dutch children listening times in the PP-V structure. Furthermore, co-activation and subsequent inhibition of Turkish can explain the less 'Turkish-like' behaviour during Dutch overt pronoun processing. In other words, CLISP can account for the online effects of cross-linguistic influence we observed in two different tasks (i.e., self-paced listening and eye-tracking), in two different populations (i.e., German-Dutch and Turkish-Dutch children), and for two different language properties (i.e., word order and pronoun resolution). We now turn to surface overlap and language dominance.

Surface overlap

Recall that situations of *partial overlap* are assumed to be more vulnerable to cross-linguistic influence than other overlap situations (i.e., *complete* and *no overlap*; e.g., Hulk & Müller, 2000; Foroodi-Nejad & Paradis, 2009; Unsworth, 2003). To be more precise, cross-linguistic influence might be most likely to occur in a situation in which the tested language has two options for a morphosyntactic property and the other language has only one option. This prediction follows from CLISP.

First of all, when there is optionality in one language, children will sometimes have to activate and select option 1 and sometimes option 2. This holds for the long passive structure in Dutch: children will sometimes hear or use the V-PP structure and will sometimes hear or use the PP-V structure (see Figure 6.1). Second, when there is no optionality in a language, children will always activate and select the same option. In order to illustrate the role of surface overlap, we assume that this holds for the long passive structure in German: only the PP-V structure is possible. Hence, a syntactic representation in a language with only one option will by definition be processed more often than a syntactic representation in a language with multiple options. Priming studies show that the more often a structure is activated and selected over time, the more likely it is to become activated and selected in the future (e.g., Bernolet, Hartsuiker, & Pickering, 2007; Kootstra & Doedens, 2016). This implies that the structure in the language with only one option is more likely to become activated and selected over time than in the language with multiple options. We also assume that combinatorial nodes will become more strongly connected to lemmas from the language with only one option than from the language with two options (Kootstra & Doedens, 2016). This is schematically represented by the thickness of the connections between the PP-V node and *worden* and *werden* in Figure 6.5.

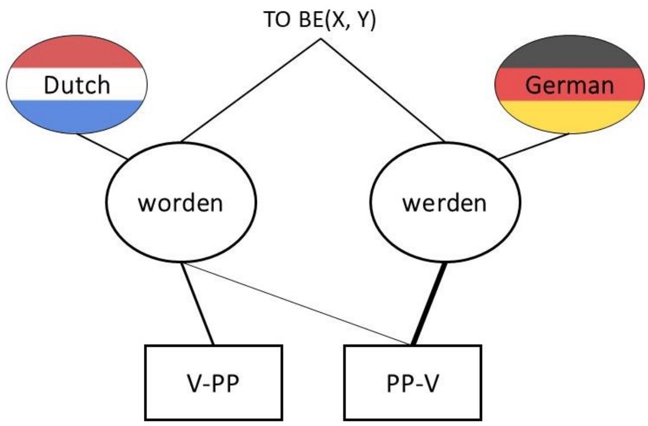


Figure 6.5. Schematic representation of the Dutch passive auxiliary *worden* and the German passive auxiliary *werden* at the lemma level.

In contrast, the situation is different with *complete* and *no overlap*. When there is *complete overlap*, both bilingual children's languages have only one and the same morphosyntactic option available in both of their languages. This is the case for the V2 structure in Dutch and German. As a consequence, children in principle process the option equally often in both of their languages. Hence, the strength of the connections between the combinatorial node and lemmas are similar between languages. This is schematically represented in Figure 6.6.

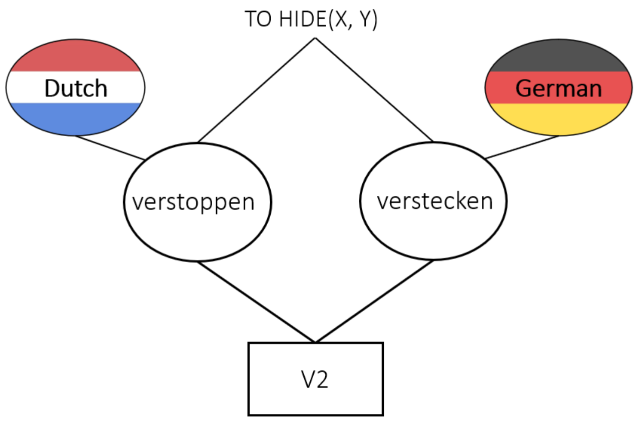


Figure 6.6. Schematic representation of the Dutch verb *verstoppen* and the German verb *verstecken* (to hide) at the lemma level.

Hence, the crucial difference between a situation of *partial* and *complete overlap* is that a combinatorial node in the first situation is stronger connected to lemmas from one language than from the other. However, in a situation of *complete overlap* connection strength is similar between languages. As a result, CLISP predicts stronger relative co-activation and, in turn, inhibition of the language not in use in a situation of *partial overlap* than in a situation of *complete overlap*. This prediction was corroborated in Chapter 3. Slowdown effects in the German-Dutch children were more pronounced in the PP-V structure than the V2 structure.

Finally, when there is *no overlap*, one language has one syntactic representation for a morphosyntactic property and the other language has another representation (e.g., Nicoladis, 2006, 2012; Nicoladis et al., 2010). This is the case for verb placement in Dutch and English. Dutch has V2 order and English has V3 order. As a consequence, in such a situation co-activation of an English lemma only occurs through the shared semantic representation of the Dutch and English lemma and not through a shared combinatorial node (e.g., Nicoladis, 2006; 2012; Nicoladis et al., 2010; Figure 6.7). Furthermore, whilst an ungrammatical V3 order in Dutch may activate the V3 node of an English lemma to some extent, we expect this activation to be weak. This is because Dutch lemmas for verbs are not connected to a V3 node. As a result, co-activation of the syntactic representation in a situation of *no overlap* may frequently be too small to result in (visible) inhibition effects during sentence processing. Indeed, we found no significant slowdown effects during the processing of Dutch V2 and ungrammatical V3 structures in English-Dutch children.

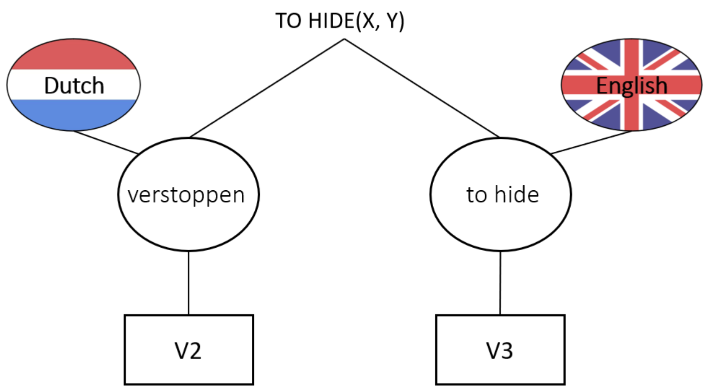


Figure 6.7. Schematic representation of the Dutch verb *verstoppen* and the English verb *to hide* at the lemma level.

Whilst the results from Chapter 3 generally support the predictions from CLISP, there are two issues that need to be addressed. First, our initial assumption that German-Dutch children do not have a V-PP representation in German was too strong. Instead, extraposition is possible in (spoken) German (e.g., Betz, 2008; Dürscheid, 2012; Haider, 2010), therefore allowing the V-PP structure to some extent. Indeed, we observed slowdown effects too when German-Dutch children processed the V-PP structure in Dutch. Consequently, the long passive structure may not have been the best testing ground for effects of surface overlap in bilingual processing. Having said that, we could still argue that the long passive structure constitutes a situation of *partial overlap* between Dutch and German due to differences in frequency of occurrence of the PP-V structures in the two languages. Second, we found no evidence of cross-linguistic influence in English-Dutch children's listening times, irrespective of the type of overlap between Dutch and English structures. This we attributed to insufficient co-activation of English in general. In other words, we cannot be certain that the absence of cross-linguistic influence with V2 and V3 structures (*no overlap*) was caused by the lack of surface overlap or by a general lack of co-activation of English. We therefore call for future studies that further investigate online effects of cross-linguistic influence in different overlap situations.

Language dominance

So far, we implicitly assumed in our discussion of the sentence processing model that bilingual children have a balanced proficiency in and exposure to both their languages. However, bilingual children are typically dominant in one of their languages (e.g., Grosjean, 1982). Cross-linguistic influence has been argued to become stronger with increased dominance in the language not in use (e.g., Bosch & Unsworth, 2020; Foroodi-Nejad & Paradis, 2009). We will now explore two reasons for such an effect of language dominance based on the sentence processing model.

First, in order to develop abstract syntactic representations in a language, and in order for representations to become shared across languages, bilinguals need to reach a certain level of proficiency in their languages (e.g., Hartsuiker & Bernolet, 2017). Hence, only with sufficient proficiency in Dutch and German, for instance, can German-Dutch children develop a shared PP-V node. When nodes are not shared, the sentence processing model predicts only indirect co-activation through semantic representations of the PP-V structure in German during Dutch sentence processing (see Figure 6.8). This is similar to the situation for the V2 and V3 structures in Dutch and English depicted in Figure 6.7. As a consequence, activation of the Dutch PP-V node will not increase the co-activation of German lemmas.

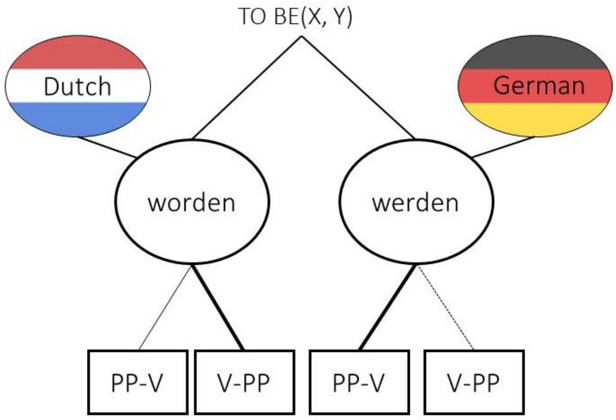


Figure 6.8. Schematic representation of the Dutch passive auxiliary *worden* and the German passive auxiliary *werden* (to be) with separate combinatorial nodes.

Second, the more dominant bilinguals are in one language, the stronger lemmas and their combinatorial nodes in this language are linked and become activated during sentence processing (e.g., Kootstra & Doedens, 2016; Kroll & Stewart, 1994; Schoonbaert et al., 2007). This is because children have more experience with this language. In other words, the more dominant bilingual children are in the language not in use, the more co-activation this language receives during sentence processing of another language. As a result, bilingual children have to allocate more processing resources to suppress co-activation. This situation is schematically depicted in Figure 6.9.

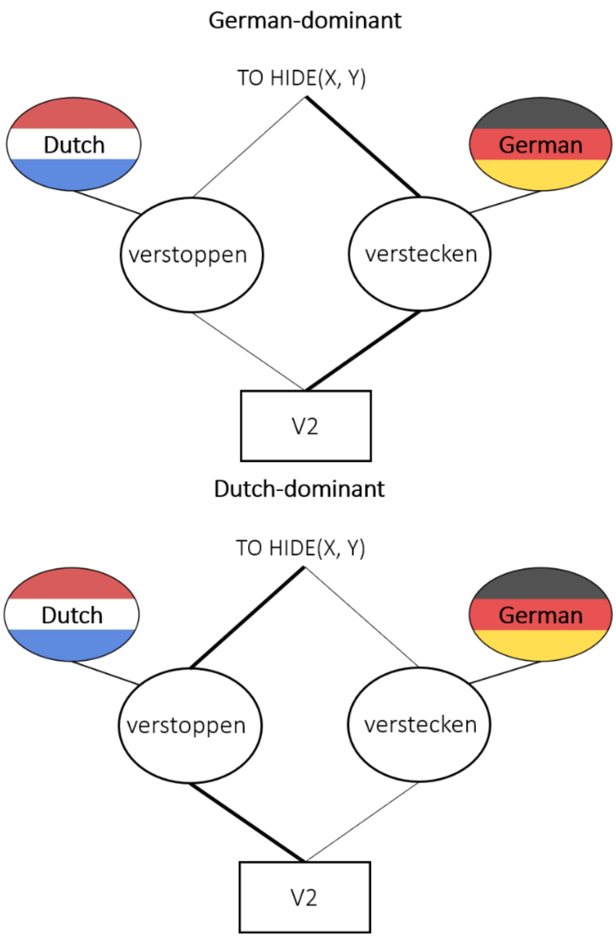


Figure 6.9. Schematic representation of the Dutch verb *verstoppen* and the German verb *verstecken* at the lemma level. The top figure depicts a situation in which a child is dominant in German. The bottom figure depicts a situation in which a child is dominant in Dutch.

The top figure represents a situation in which the child is German-dominant. The connection between the German lemma *verstecken* and the V2 node is stronger than between the Dutch lemma *verstoppen* and the V2 node, as depicted by the thickness of the lines (e.g., Kootstra & Doedens, 2016). As a result, German lemmas will be strongly co-activated during the processing of Dutch V2 sentences. Hence, the activation of German lemmas needs to be strongly inhibited. The bottom picture represents a situation in which a German-Dutch child is dominant in Dutch. The connection between the German lemma *verstecken* and the V2 node is relatively weak.

Consequently, German lemmas will receive little co-activation during the processing of Dutch V2 sentences. Hence, the activation of German lemmas need only be weakly inhibited. In other words, cross-linguistic influence from German during Dutch sentence processing is stronger in the first than in the second situation.

The results from both our online studies with bilingual children corroborate the predictions from CLISP. With regard to the self-paced listening task with German-Dutch bilingual children, the more dominant children were in German, the more they slowed down when processing Dutch PP-V and V2 structures. With regard to the eye-tracking study with Turkish-Dutch bilingual children, the more dominant children were in Turkish, the less they looked at the Turkish-preferred non-topic referent when processing Dutch overt pronoun structures. These effects were significant when language dominance was operationalized as children's relative language proficiency. Similar non-significant trends were observed for children's relative current and cumulative language input. There was one exception: cumulative exposure predicted online cross-linguistic influence from German to Dutch in the V2 structure.

The observation that only children's relative proficiency and not their relative input significantly predicted the strength of cross-linguistic influence for the long passive structure and overt pronouns is in line with our first explanation of effects of language dominance described above. The long passive structure and the discourse properties of pronouns are acquired relatively late (e.g., Armon-Lotem et al., 2016; Arnold, Brown-Schmidt, & Trueswell, 2007; Bartke, 2004; Järvikivi, Pyykkönen-Klauck, Schimke, Colonna, & Hemforth, 2014; Song & Fisher, 2005; Verrips, 1996). Consequently, (some of) the Dutch-dominant children in our studies might not have acquired abstract syntactic and semantic representations for long passives and pronouns in (one of) their languages, and, consequently might not have developed shared combinatorial nodes (e.g., Hartsuiker & Bernolet, 2017). Hence, children's relative proficiency level might have reflected the availability of shared combinatorial nodes (in addition to the amount of co-activation).

In contrast, the V2 structure in Dutch and German is acquired early (e.g., Blom, 2003; Clahsen & Penke, 1992; Poeppel & Wexler, 1993; Wijnen & Verrips, 1998) and, therefore, likely has developed in a shared combinatorial node with stable connections to Dutch and German lemmas for most children tested. As a result, individual differences in co-activation of German during the processing of V2 structures are only the result of differences in the strength of connections between semantic representations, combinatorial nodes and lemmas, and not of the presence or absence of shared combinatorial nodes. Both relative proficiency and language input can be argued to influence these connection strengths, explaining why relative proficiency *and* cumulative input predicted the strength of cross-linguistic influence with V2 structures. In other words, whilst both relative proficiency and language input might be good predictors of online cross-linguistic influence in older bilingual children for morphosyntactic properties that are acquired early, relative proficiency might be a better predictor of online cross-linguistic influence than input measures in young bilingual children and for morphosyntactic properties that are acquired late.

In short, in the processing model effects of language dominance are a logical consequence of the strength of connections between lemma nodes (e.g., Kootstra & Doedens, 2016; Kroll & Stewart, 1994; Schoonbaert et al., 2007). In our studies, we were able to investigate dominance effects ranging from children being dominant in the language tested (Dutch) to children being more or less balanced in both of their languages. How dominance affects online cross-linguistic influence in children that are dominant in the language not in use requires further research. For Dutch self-paced listening experiments with German or English as dominant languages, we would make two predictions: (i) German-dominant children should show larger slowdown effects than observed for the children in our study; and (ii) English-dominant children may show slowdown effects in the V-PP and V3 structures not observed for the children in our study. For the eye-tracking study, we predict the opposite effect to occur in Turkish-dominant children than observed for the balanced bilinguals: Turkish-dominant children might have insufficient processing resources to sufficiently inhibit the (stronger) co-activation of the Turkish non-topic preference for Dutch overt pronouns. Consequently, children should fixate more on the non-topic referent than balanced bilinguals. Future studies are necessary to test these predictions.

Effects of cross-linguistic influence in sentence production and offline comprehension

So far, we explained online effects of cross-linguistic influence. However, the CLISP model can also account for cross-linguistic influence in children's sentence production and offline comprehension. There are two mechanisms that explain how language co-activation and inhibition *during* sentence processing can result in cross-linguistic influence in bilingual children's responses: (i) the inability to resolve inhibition effects; and (ii) the lack of working memory resources. Cross-linguistic influence arises due to inhibition effects when a child suppresses a semantic representation or combinatorial node shared between her languages so strongly during sentence processing that this representation is not available for selection by the time she produces an utterance or makes an offline decision. In Chapter 5, we argued that this sometimes happened in children's offline referent choices for the overt pronoun. To be more precise, the stronger children suppressed the co-activation of the Turkish pronoun *o* and its preferred interpretation during sentence processing, the less likely it was that they were able to access this interpretation later on (see Figure 6.10). Such effects are similar to other effects of cross-linguistic influence we observed in our meta-analysis (Chapter 2) that we deemed inconsistent with cross-linguistic influence (e.g., Mykhaylyk & Ytterstad, 2017; Nicoladis, 2006) and effects of overcorrection in simultaneous bilingual adults (e.g., Anderssen et al., 2018; Kupisch, 2014).

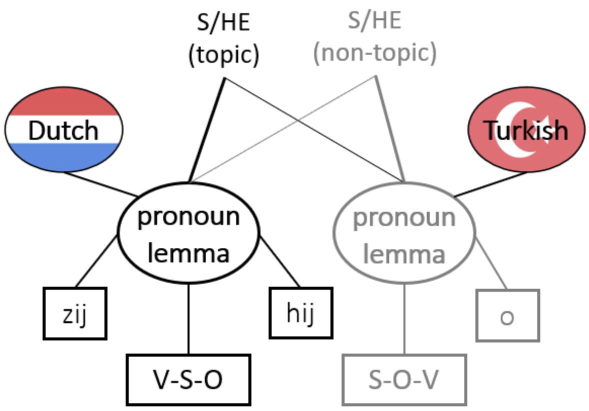


Figure 6.10. Schematic representation of the Dutch and Turkish pronoun lemma, their semantic representations, combinatorial nodes and forms.

Cross-linguistic influence can also occur when the co-activation of the language not in use is so strong that children's maximum working memory capacity is reached. In this situation, representations that are not sufficiently activated will be lost from working memory (e.g., Just & Carpenter, 1992). Consequently, if at this moment a semantic representation or combinatorial node from the language not in use is the most strongly activated one, it can become selected over a representation from the language in use. For instance, in the context of our overt pronoun experiment, working memory capacity limitations could result in an increase in non-topic interpretations offline (especially in Turkish-dominant children). Such effects correspond with cross-linguistic influence typically attested in children's sentence production and offline comprehension (e.g., Argyri & Sorace, 2007; Bosch & Unsworth, 2020; Foroodi-Nejad & Paradis, 2009; also see the meta-analysis in Chapter 2). Note that such an account also explains why bilingual children sometimes use or accept a structure from their one language that is not available in their other language (in a situation of *no overlap*; e.g., Nicoladis, 2006; Nicoladis & Gavrila, 2015), such as ungrammatical V2 orders in English (e.g., Bosch & Unsworth, 2020) or V3 orders in Dutch or German (e.g., Döpke, 1998; Unsworth, 2016).

Given that offline tasks – and we believe elicited production tasks as well – place a burden on children's working memory capacity (e.g., Marinis, 2010), cross-linguistic influence offline may be more likely to occur due to working memory limitations than due to too strong inhibition effects. Indeed, the results of our meta-analysis are in line with this expectation (Chapter 2): 24 observations of cross-linguistic influence could be interpreted as resulting from inhibition effects versus 73 observations that could be interpreted as resulting from working memory capacity limitations.

Furthermore, in our view, an important factor that determines whether cross-linguistic influence surfaces in children's sentence production or offline comprehension due to inhibition effects or working memory capacity limitations is the strength of co-activation of the language not in use. For inhibition effects to become present in children's linguistic choices, co-activation has to be strong enough for online inhibition effects to last. At the same time, if co-activation is so strong that it cannot be fully suppressed, children might run out of working memory capacity and select a co-activated semantic representation or combinatorial node from the language not in use. In our overt pronoun study, we observed that offline inhibition effects increased – although not significantly so – the more balanced children were in their Dutch and Turkish. We predict that with increased dominance in Turkish, these effects offline (and online) will reverse. In other words, for more balanced bilingual children, we predict offline inhibition effects of the language not in use. For unbalanced bilingual children, we predict the overuse of representations from their dominant language.

Finally, CLISP also explains why online effects of cross-linguistic influence were more pronounced than offline effects in the overt pronoun experiment. Following the explanation we gave for offline cross-linguistic influence, online inhibition effects only become visible offline when inhibition is too strong to resolve. When inhibition is only experienced temporally during sentence processing, cross-linguistic influence will only surface online and not offline. In other words, if cross-linguistic influence offline is caused by language co-activation online, inhibition effects offline are a weaker reflection of inhibition effects online.

Online cross-linguistic influence in adulthood

Finally, what is left to account for are effects in simultaneous bilingual adults. Recall that in this thesis we found similar but less pronounced slowdown effects in our self-paced listening task with adults and adolescents than with children. Furthermore, previous online and offline studies suggest that cross-linguistic influence is more likely to become visible as overcorrection in adults than in children (e.g., Kupisch, 2014; Kupisch & Barton, 2013; Schmitz, Di Venanzio, & Scherger, 2016). Both of these observations follow from the CLISP model when we assume that simultaneous bilingual adults have more working memory resources available for sentence processing and/or make more efficient use of these resources than children (e.g., Chi, 1978; Dempster, 1981; Kail, 1991; also see Schneider, 2015 for a review of the literature). As a result, adults are more likely to successfully inhibit language co-activation (e.g., Linck, Hoshino, & Kroll, 2008). This has two consequences.

First, the chance that adults run out of working memory resources while they have activated representations from both of their languages is smaller than in children. Therefore, the erroneous selection of a representation from the non-target language is less likely in the adult group. Hence, CLISP predicts that the overuse of a morphosyntactic property or an interpretation from the language not in use happens less often in adults than in children.[5]

Second, because adults can make more efficient use of their working memory resources, we believe them to be more likely to successfully inhibit language co-activation than children, even when co-activation is strong. However, as we argued before, such strong inhibition can result in the – temporal – unavailability of a particular semantic representation or combinatorial node shared between bilinguals' languages. Consequently, the ability of adult bilinguals to recruit more working memory resources than children results in more successful inhibition of the language not in use, and, as a result, more overcorrection.

---

[5] Note that this reasoning also suggests that cross-linguistic influence in the form of overuse of morphosyntactic properties and interpretations decreases in older children. We did not observe a significant effect of age on cross-linguistic influence in the meta-analysis (Chapter 2), however. Other factors might compensate for an effect of age, such as more stable shared syntactic representations in older children (increasing co-activation from the language not in use compared to younger children) and differences in the age of acquisition of morphosyntactic properties tested in younger and older children (discussed in Chapter 2).

Limitations and perspectives for future research
One strength of the CLISP model is that it does not only account for most of our online and offline findings in bilingual children and adults, as we have illustrated in the preceding section, but that it also allows us to make testable predictions about bilingual sentence processing. At the same time, there are a number of limitations to our work that need to be addressed. First of all, CLISP is based on a small number of online studies investigating three language combinations and two morphosyntactic properties (word order and pronoun resolution). Thus, its generalisability still needs to be assessed. Second, as mentioned above already, we have not tested a group of bilingual children or adults who were dominant in the language that was not tested in our experiment. Nevertheless, the model allowed us to formulate a number of predictions for the online behaviour of this group. Third, to the best of our knowledge, our eye-tracking study is the only study to date that directly compares cross-linguistic influence in bilingual children's online and offline sentence processing. Hence, it is unknown how online and offline cross-linguistic influence relate for other online (e.g., self-paced listening) and offline (e.g., elicited production and acceptability judgement) tasks. In sum, future studies are necessary to test whether the assumptions and predictions of the model generalize to different populations, for different morphosyntactic properties and for different task types.

Furthermore, future studies should explore factors that might affect language co-activation that we were unable to assess (in detail). Three examples of such factors are lexical overlap, working memory and inhibitory control. With regard to lexical overlap, the results of Chapter 3 suggest that increased lexical overlap between languages increases co-activation and, therefore, visible effects of cross-linguistic influence in children's speech production and offline comprehension. However, the self-paced listening experiment was not constructed to directly test for lexical effects within the language combinations tested. Future studies should therefore manipulate the amount of lexical overlap between sentences in experiments (e.g., Hopp, 2017). In addition, we predict that effects of cross-linguistic influence online and in offline comprehension and production should decrease the better a bilingual child's working memory capacity and inhibitory control skills. In Chapter 3, we indeed found that working memory predicted children's listening times. However, we did not test for effects of working memory and inhibitory control on effects of cross-linguistic influence.

As the above discussion has shown, we believe that the sentence processing model presented in this chapter offers a fruitful theoretical framework for future online and offline studies on cross-linguistic influence that set out to investigate the topics in the above in bilingual children and adults.

6.4 Implications

What do the results of this thesis tell us about cross-linguistic influence in simultaneous bilinguals? First and foremost, this thesis confirms the well-known observation that a bilingual child is not two monolinguals in one (e.g., Grosjean, 1989). The languages of bilingual children are interacting with each other, even when a child is using only one of his or her two languages (i.e., is in a monolingual mode). This means that this interaction will be visible during normal day-to-day activities, for instance at home or at school. This interaction sometimes results in children overusing, for example, the word order of their one language in their other. At the same time, this thesis also shows that cross-linguistic influence may often be invisible: parents and teachers may not notice that it takes places because it does not necessarily surface in the way in which bilingual children speak and understand their two languages. Instead, language co-activation may only result in subtle effects *during* sentence processing. Third, not every child and not every morphosyntactic property is equally sensitive to cross-linguistic influence. In the following, we discuss three factors that can impact the presence and strength of cross-linguistic influence: similarities between morphosyntactic properties, language dominance and age.

The results in this thesis suggest that when a bilingual child hears a morphosyntactic property in one language, cross-linguistic influence is most likely to occur when their other language has a similar property. This holds, for example, for word order: similar word orders between languages are likely to activate the language not in use. Chances of cross-linguistic influence are further increased when the word order in question is more frequent in the language not in use than in the language in use. If our perspective on the role of overlapping morphosyntactic properties is correct, children who acquire two languages that share many morphosyntactic properties may experience more cross-linguistic influence than children with very different languages.

With regard to language dominance, results from this thesis as well as from previous studies show that cross-linguistic influence becomes stronger, and is therefore more likely to become visible, the more dominant children are in the language not in use. In other words, a Turkish-Dutch bilingual child growing up and going to school in, for example, the Netherlands, is likely to be Dutch-dominant and, therefore, likely to experience stronger cross-linguistic influence from Dutch into Turkish than from Turkish into Dutch. At the same time, Turkish-Dutch bilingual peers may still differ from each other: the more proficient children are in Turkish, the stronger the influence from Turkish to Dutch. Hence, the situation with respect to cross-linguistic influence will be different for a Turkish-Dutch child in the Netherlands who always speaks Turkish outside of school (for example, at home and with friends) and for a Turkish-Dutch child who only speaks Turkish a few hours a week with her grandmother. The strength of cross-linguistic influence experienced thus depends on children's dominance profile, and again, even when it occurs, it might not be visible to parents and teachers.

With regard to age, our results with simultaneous bilingual adults and adolescents suggest two things. On the one hand, cross-linguistic influence persists with age: even bilingual adults experience subtle effects of language co-activation. On the other hand, cross-linguistic influence may become less strong as bilinguals get older. In our view, this is because adults have more cognitive resources available and can therefore deal more efficiently with language co-activation. If our assumption is correct, a young child might be more strongly influenced by morphosyntactic properties of her other language than older children and adults. Consequently, younger children may more often show visible cross-linguistic influence than older children and adults, for instance, by using the word order in one language from their other language.

In sum, this thesis shows that cross-linguistic influence is part and parcel of being bilingual. Whether it occurs and how strong it is depends on various factors. Importantly, cross-linguistic influence occurs more often than we may actually realize. At the same time, our studies show that cross-linguistic influence in children and adults is subtle and often only visible with detailed experimental techniques. This observation shows that acquiring two languages in parallel results in effective language parsing and that bilingual children grow into very skilled language users.

6.5 Conclusions

In this thesis, we have investigated effects of cross-linguistic influence during and after sentence processing in simultaneous bilingual children, adolescents and adults. To our knowledge, we were the first to observe direct evidence of cross-linguistic influence online in bilingual children. Such influence occurred in the form of inhibition effects. Furthermore, children's online behaviour was affected by the type of surface overlap and the amount of lexical overlap between children's languages and children's dominance profile, in line with existing studies on children's sentence production and online comprehension.

We found no evidence for negative effects during sentence processing due to bilingualism in general.

Whilst online patterns from our eye-tracking study were reflected in an offline comprehension task, the offline effects were less pronounced and did not reach significance. This finding suggest that online tasks are more direct measures of cross-linguistic influence than offline tasks. Finally, some of the online effects observed in the self-paced listening task for children were observed for bilingual adolescents and adults as well. These last two findings suggests that similar mechanisms result in cross-linguistic influence online and offline and during different stages of bilingual acquisition.

One important result of this thesis is the comprehensive model that we developed (CLISP) on the basis of the collected empirical data and of existing accounts of bilingual sentence processing. This new model did not only account for our results, but also allowed us to explain – sometimes contrasting – effects of cross-linguistic influence from existing production and offline comprehension tasks. Because it also makes detailed predictions, the model can serve as a tool for future research to explore sentence processing and cross-linguistic influence in bilingual children and adults.

References

Alemán Bañón, J., Fiorentino, R., & Gabriele, A. (2014). Morphosyntactic processing in advanced second language (L2) learners: An event-related potential investigation of the effects of L1–L2 similarity and structural distance. *Second Language Research*, *30*(3), 275–306. https://doi.org/10.1177/0267658313515671

Alemán Bañón, J., Fiorentino, R., & Gabriele, A. (2018). Using event-related potentials to track morphosyntactic development in second language learners: The processing of number and gender agreement in Spanish. *PloS ONE*, *13*(7), 1–35. https://doi.org/10.1371/journal.pone.0200791

Alloway, T. P. (2012). *Alloway Working Memory Assessment 2 (AWMA-2)*. London: Pearson.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*(3), 191–238. https://doi.org/10.1016/0010-0277(88)90020-0

Anderssen, M., Lundquist, B., & Westergaard, M. (2018). Cross-linguistic similarities and differences in bilingual acquisition and attrition: Possessives and double definiteness in Norwegian heritage language. *Bilingualism: Language and Cognition*, *21*(4), 748–764. https://doi.org/10.1017/S1366728918000330

Andersson, A., Sayehli, S., & Gullberg, M. (2019). Language background affects online word order processing in a second language but not offline. *Bilingualism: Language and Cognition*, *22*(4), 802–825. https://doi.org/10.1017/S1366728918000573

Argyri, E., & Sorace, A. (2007). Crosslinguistic influence and language dominance in older bilingual children. *Bilingualism: Language and Cognition*, *10*(1), 79–99. https://doi.org/10.1017/S1366728906002835

Ariel, M. (1994). Interpreting anaphoric expressions: a cognitive versus a pragmatic approach. *Journal of Linguistics*, *30*(1), 3–42. https://doi.org/10.1017/S0022226700016170

Ariel, M. (2014). *Accessing noun-phrase antecedents* (2nd ed.). Routledge. https://doi.org/10.4324/9781315857473

Armon-Lotem, S., Haman, E., Jensen de López, K., Smoczynska, M., Yatsushiro, K., Szczerbinski, M., … van der Lely, H. (2016). A large-scale cross-linguistic investigation of the acquisition of passive. *Language Acquisition*, *23*(1), 27–56. https://doi.org/10.1080/10489223.2015.1047095

Arnold, J. E., Brown-Schmidt, S., & Trueswell, J. (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes*, *22*(4), 527–565. https://doi.org/10.1080/01690960600845950

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition*, *76*(1), B13–B26. https://doi.org/10.1016/S0010-0277(00)00073-1

Austin, J. (2007). Grammatical interference and the acquisition of ergative case in bilingual children learning Basque and Spanish. *Bilingualism: Language and Cognition*, *10*(3), 315–331. https://doi.org/10.1017/S1366728907003094

Austin, J. (2009). Delay, interference and bilingual development: The acquisition of verbal morphology in children learning Basque and Spanish. *International Journal of Bilingualism*, *13*(4), 447–479. https://doi.org/10.1177/1367006909353234

Azar, Z., & Özyürek, A. (2015). Discourse management: Reference tracking in speech and gesture in Turkish narratives. *Dutch Journal of Applied Linguistics*, *4*(2), 222–240. https://doi.org/10.1075/dujal.4.2.06aza

Azar, Z., Özyürek, A., & Backus, A. (2020). Turkish-Dutch bilinguals maintain language-specific reference tracking strategies in elicited narratives. *International Journal of Bilingualism*, *24*(2), 376–409. https://doi.org/10.1177/1367006919838375

Bacchini, S., Boland, T., Hulsbeek, M., Pot, H., & Smits, M. (2005). *Duizend-en-een-woorden De allereerste Nederlandse woorden voor anderstalige peuters en kleuters*. Stichting Leerplanontwikkeling.

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

Bartke, S. (2004). Passives in German children with Williams syndrome. In S. Bartke & J. Siegmüller (Eds.), *Williams Syndrome across Languages* (pp. 345–370). John Benjamins. https://doi.org/10.1075/lald.36

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lm4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bergmann, C., Tsuji, S., Piccinini, P. E., Lewis, M. L., Braginsky, M., Frank, M. C., & Cristia, A. (2018). Promoting replicability in developmental research through meta-analyses: Insights from language acquisition research. *Child Development*, *89*(6), 1996–2009. https://doi.org/10.1111/cdev.13079

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2007). Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(5), 931–949. https://doi.org/10.1037/0278-7393.33.5.931

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2009). Persistence of emphasis in language production: A cross-linguistic approach. *Cognition*, *112*(2), 300–317. https://doi.org/10.1016/j.cognition.2009.05.013

Betz, E. (2008). *Grammar and Interaction: Pivots in German Conversation*. John Benjamins. https://doi.org/10.1075/sidag.21

Bever, T. G. (1970). The cognitive basis for linguistic structures. In R. Hayes (Ed.), *Cognition and Language Development* (pp. 279–362). Wiley & Sons.

Blom, E. (2003). *From root infinitive to finite sentence: The acquisition of verbal inflections and auxiliaries* [doctoral dissertation, Utrecht University]. LOT. https://www.lotpublications.nl/from-root-infinitive-to-finite-sentence-from-root-infinitive-to-finite-sentence

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., & Leseman, P. (2014). The benefits of being bilingual: Working memory in bilingual Turkish–Dutch children. *Journal of Experimental Child Psychology*, *128*, 105–119. https://doi.org/10.1016/j.jecp.2014.06.007

Blom, E., & Vasić, N. (2011). The production and processing of determiner–noun agreement in child L2 Dutch. *Linguistic Approaches to Bilingualism*, *1*(3), 265–290. https://doi.org/10.1075/lab.1.3.03blo

Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer* (version 6.0.37) [Computer software]. http://www.praat.org/

Bohnacker, U., Lindgren, J., & Öztekin, B. (2016). Turkish- and German-speaking bilingual 4-to-6-year-olds living in Sweden: Effects of age, SES and home language input on vocabulary production. *Journal of Home Language Research*, *1*, 17–41. https://doi.org/10.16993/jhlr.26

Bonfieni, M., Branigan, H. P., Pickering, M. J., & Sorace, A. (2019). Language experience modulates bilingual language control: The effect of proficiency, age of acquisition, and exposure on language switching. *Acta Psychologica*, *193*, 160–170. https://doi.org/10.1016/j.actpsy.2018.11.004

Booth, J. R., Mac Whinney, B., & Harasaki, Y. (2000). Developmental differences in visual and auditory processing of complex sentences. *Child Development*, *71*(4), 981–1003. https://doi.org/10.1111/1467-8624.00203

Borenstein, M., Hedges, L. V, Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons. https://doi.org/10.1002/9780470743386

Bosch, J. E., & Unsworth, S. (2020). Cross-linguistic influence in word order: effects of age, dominance and surface overlap. *Linguistic Approaches to Bilingualism*, 1–34. https://doi.org/10.1075/lab.18103.bos

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*(1), 1–38. https://doi.org/10.5334/joc.72

Cardinaletti, A., & Starke, M. (1999). The typology of structural deficiency: A case study of the three classes of pronouns. In H. van Riemsdijk (Ed.), *Clitics in the Languages of Europe* (pp. 145–234). De Gruyter Mouton. https://doi.org/10.1515/9783110804010.145

Carminati, M. N. (2002). *The processing of Italian subject pronouns* [Doctoral disseration, University of Massachusetts]. Proquest. https://scholarworks.umass.edu/dissertations/AAI3039345

Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S.

Siegler (Ed.), *Children's thinking: What develops?* (pp. 73–96). Lawrence Erlbaum Associates.

Cho, S.-J., Brown-Schmidt, S., & Lee, W. (2018). Autoregressive generalized linear mixed effect models with crossed random effects: An application to intensive binary time series eye-tracking data. *Psychometrika*, *83*(3), 751–771. https://doi.org/10.1007/s11336-018-9604-2

Chondrogianni, V., & Marinis, T. (2012). Production and processing asymmetries in the acquisition of tense morphology by sequential bilingual children. *Bilingualism: Language and Cognition*, *15*(1), 5–21. https://doi.org/10.1017/S1366728911000368

Chondrogianni, V., Marinis, T., Edwards, S., & Blom, E. (2015). Production and on-line comprehension of definite articles and clitic pronouns by Greek sequential bilingual children and monolingual children with specific language impairment. *Applied Psycholinguistics*, *36*(5), 1155–1191. https://doi.org/10.1017/S0142716414000101

Chondrogianni, V., Vasić, N., Marinis, T., & Blom, E. (2015). Production and on-line comprehension of definiteness in English and Dutch by monolingual and sequential bilingual children. *Second Language Research*, *31*(3), 309–341. https://doi.org/10.1177/0267658314564461

Chondrogianni, V. (2008). Comparing child and adult L2 acquisition of the Greek DP. In B. Haznedar & E. Gavruseva (Eds.), *Current trends in child second language acquisition: A generative perspective* (pp. 97–142). John Benjamins.

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

Clahsen, H., & Penke, M. (1992). The acquisition of agreement morphology and its syntactic consequences: New evidence on German child language from the Simone corpus. In J. Meisel (Ed.), *The Acquisition of Verb Placement* (pp. 181–223). Kluwer.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Contemori, C., Asiri, O., & Perea Irigoyen, E. D. (2019). Anaphora resolution in L2 English: An analysis of discourse complexity and cross-linguistic interference. *Studies in Second Language Acquisition*, *41*(5), 971–998. https://doi.org/10.1017/S0272263119000111

Contemori, C., & Dussias, P. E. (2020). The processing of subject pronouns in highly proficient L2 speakers of English. *Glossa: A Journal of General Linguistics*, *5*(1), 1–19. https://doi.org/10.5334/gjgl.972

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, *123*(4), 354–373. https://doi.org/10.1037/0096-3445.123.4.354

Costa, A. (2004). Speech Production in Bilinguals. In T. K. Bhatia & W. C. Ritchie (Eds.), *The Handbook of Bilingualism* (pp. 201–223). Blackwell.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491–511. https://doi.org/10.1016/j.jml.2004.02.002

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(5), 1057–1074. https://doi.org/10.1037/0278-7393.32.5.1057

Cunnings, I. (2017). Parsing and Working Memory in Bilingual Sentence Processing. *Bilingualism: Language and Cognition*, *20*(4), 659–678. https://doi.org/10.1017/S1366728916000675

Cunnings, I., Fotiadou, G., & Tsimpli, I. (2017). Anaphora resolution and reanalysis during L2 sentence processing: Evidence from the visual world paradigm. *Studies in Second Language Acquisition*, *39*(4), 621–652. https://doi.org/10.1017/s0272263116000292

Cuza, A., & Pérez-Tattam, R. (2016). Grammatical gender selection and phrasal word order in child heritage Spanish: A feature re-assembly approach. *Bilingualism: Language and Cognition*, *19*(1), 50–68. https://doi.org/https://doi.org/10.1017/S1366728914000893

Daskalaki, E., Chondrogianni, V., Blom, E., Argyri, F., & Paradis, J. (2019). Input effects across domains: The case of Greek subjects in child heritage language. *Second Language Research*, *35*(3), 421–445.

https://doi.org/10.1177/0267658318787231

De Houwer, A. (2011). Language input environments and language development in bilingual acquisition. *Applied Linguistics Review*, *2*(2), 221–240.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, *23*(4), 517–542. https://doi.org/10.1207/s15516709cog2304_6

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin*, *89*(1), 63–100. https://doi.org/10.1037/0033-2909.89.1.63

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*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., van Jaarsveld, H., & ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, *1*(1), 51–66.

Dijkstra, T., Wahl, A., Buytenhuijs, F., van Halem, N., Al-Jibouri, Z., de Korte, M., & Rekké, S. (2019). Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679. https://doi.org/10.1017/S1366728918000287

Döpke, S. (1998). Competing language structures: the acquisition of verb placement by bilingual German-English children. *Journal of Child Language*, *25*(3), 555–584. https://doi.org/10.1017/S0305000998003584

Dürscheid, C. (2012). *Syntax. Grundlagen und Theorien*. VandenHoeck & Ruprecht.

Edwards, H. T., & Kirkpatrick, A. G. (1999). Metalinguistic awareness in children: A developmental progression. *Journal of Psycholinguistic Research*, *28*(4), 313–329. https://doi.org/10.1023/A:1023275214000

Egger, E., Hulk, A., & Tsimpli, I. M. (2018). Crosslinguistic influence in the discovery of gender: The case of Greek–Dutch bilingual children. *Bilingualism: Language and Cognition*, *21*(4), 694–709. https://doi.org/10.1017/S1366728917000207

Enç, M. (1986). Topic Switching and Pronominal Subjects in Turkish. In D. I. Slobin & K. Zimmer (Eds.), *Studies in Turkish Linguistics* (pp. 195–209). John Benjamins. https://doi.org/10.1075/tsl.8.11enc

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Felser, C. (2020). Do processing resource limitations shape heritage language grammars? *Bilingualism: Language and Cognition*, *23*(1), 23–24. https://doi.org/10.1017/S1366728919000397

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, *24*(3), 453–489. https://doi.org/10.1017/S0142716403000237

Fernández-Castilla, B., Jamshidi, L., Declercq, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2020). The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behavior Research Methods*, *52*, 2031–2052. https://doi.org/10.3758/s13428-020-01373-9

Fernández-Castilla, B., Maes, M., Declercq, L., Jamshidi, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2019). A demonstration and evaluation of the use of cross-classified random-effects models for meta-analysis. *Behavior Research Methods*, *51*, 1286–1304. https://doi.org/10.3758/s13428-018-1063-2

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*(3), 348–368. https://doi.org/10.1016/0749-596X(86)90006-9

Ferreira, F., Henderson, J. M., Anes, M. D., Weeks, P. A., & McFarlane, D. K. (1996). Effects of lexical frequency and syntactic complexity in spoken-language comprehension: Evidence from the auditory moving-window technique. *Journal of Experimental Psychology: Learning Memory and*

Cognition, *22*(2), 324–335. https://doi.org/10.1037/0278-7393.22.2.324

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, *40*(4), 296–340. https://doi.org/10.1006/cogp.1999.0730

Filiaci, F. (2010). Null and overt subject biases in Spanish and Italian: A cross-linguistic comparison. In C. Borgonovo, M. Español-Echevarría, & P. Prévost (Eds.), *Selected Proceedings of the 12th Hispanic Linguistics Symposium* (pp. 171–182). Cascadilla Proceedings Project.

Foroodi-Nejad, F., & Paradis, J. (2009). Crosslinguistic transfer in the acquisition of compound words in Persian–English bilinguals. *Bilingualism: Language and Cognition*, *12*(4), 411–427. https://doi.org/10.1017/S1366728909990241

Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, *66*(1), 226–248. https://doi.org/10.1016/j.jml.2011.07.007

Fox, D., & Grodzinsky, Y. (1998). Children's passive: A view from the by-phrase. *Linguistic Inquiry*, *29*(2), 311–332. https://doi.org/10.1162/002438998553761

Gathercole, V. C. M., Laporte, N., & Thomas, E. M. (2005). Differentiation, carry-over, and the distributed characteristic in bilinguals: Structural "mixing" of the two languages? In J. Cohen, K. T. McAlister, K. Rolstad, & J. MacSwan (Eds.), *Proceedings of the 4th International Symposium on Bilingualism (ISB4)* (pp. 838–851). Cascadilla Press.

Gavarró, A. (2003). Economy and word order patterns in bilingual English-Dutch acquisition. *Bilingualism: Language and Cognition*, *6*(1), 69–79. https://doi.org/10.1017/S1366728903001044

Genesee, F., Paradis, J., & Crago, M. B. (2004). *Dual language development & disorders: A handbook on bilingualism & second language learning*. Paul Brookes.

Gillon Dowens, M., Guo, T., Guo, J., Barber, H., & Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish - Evidence from Event Related Potentials. *Neuropsychologia*, *49*(7), 1651–1659.

https://doi.org/10.1016/j.neuropsychologia.2011.02.034

Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed.), *Topic continuity in discourse: a quantitative cross-language study* (pp. 1–42). https://doi.org/10.1075/tsl.3.01giv

Goriot, C., Broersma, M., McQueen, J. M., Unsworth, S., & van Hout, R. (2018). Language balance and switching ability in children acquiring English as a second language. *Journal of Experimental Child Psychology*, *173*, 168–186. https://doi.org/10.1016/j.jecp.2018.03.019

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67–81. https://doi.org/10.1017/S1366728998000133

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Harvard University Press.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*(1), 3–15. https://doi.org/10.1016/0093-934X(89)90048-5

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, *1*(2), 131–149. https://doi.org/10.1017/S136672899800025X

Grosjean, F. (2001). The bilingual's language modes. In J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Wiley-Blackwell.

Guijarro-Fuentes, P. (2019). Cross-linguistic influence in bilingual language acquisition: Determining onset and end. *Probus*, *31*(2), 233–244. https://doi.org/10.1515/probus-2016-0010

Hacohen, A., & Schaeffer, J. (2007). Subject realization in early Hebrew/English bilingual acquisition: The role of crosslinguistic influence. *Bilingualism: Language and Cognition*, *10*(3), 333–344. https://doi.org/10.1017/S1366728907003100

Haider, H. (2010). *The syntax of German*. Cambridge University Press.

https://doi.org/10.1017/CBO9780511845314

Haman, E., Łuniewska, M., & Pomiechowska, B. (2015). Designing cross-linguistic lexical tasks (CLTs) for bilingual preschool children. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing Multilingual Children* (pp. 196–240). https://doi.org/10.21832/9781783093137-010

Hamann, C., & Abed Ibrahim, L. (2017). Methods for Identifying Specific Language Impairment in Bilingual Populations in Germany. *Frontiers in Communication*, *2*(16). https://doi.org/10.3389/fcomm.2017.00016

Hartsuiker, R. J., & Bernolet, S. (2017). The development of shared syntax in second language learning. *Bilingualism: Language and Cognition*, *20*(2), 219–234. https://doi.org/10.1017/S1366728915000164

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science*, *15*(6), 409–414. https://doi.org/10.1111/j.0956-7976.2004.00693.x

Haznedar, B. (2007). Crosslinguistic influence in Turkish-English bilingual first language acquisition: The overuse of subjects in Turkish. In A. Belikova, L. Meroni, & M. Umeda (Eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)* (pp. 124–134). Cascadilla Proceedings Project.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://doi.org/10.2307/1164588

Hedges, L. V, & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.

Hemforth, B., Konieczny, L., Scheepers, C., Colonna, S., Schimke, S., Baumann, P., & Pynte, J. (2010). Language specific preferences in anaphor resolution: Exposure or gricean maxims? In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 2218–2223). Cognitive Science Society.

Hervé, C., & Lawyer, L. (unpublished manuscript). *Bilinguals' sensitivity to specificity and genericity: evidence from implicit and explicit knowledge*. 1–52.

Hervé, C., Serratrice, L., & Corley, M. (2016). Dislocations in French–English bilingual children: An elicitation study. *Bilingualism: Language and Cognition*, *19*(5), 987–1000. https://doi.org/10.1017/S1366728915000401

Hirsch, C., & Wexler, K. (2006). Children's passives and their resulting interpretation. In K. U. Deen, J. Nomura, B. Schulz, & B. D. Schwartz (Eds.), *The proceedings of the inaugural Conference on Generative Approaches to Language Acquisition-North America (GALANA)* (pp. 125–136). University of Connecticut Occassional Papers in Linguistics.

Hopp, H. (2009). The syntax–discourse interface in near-native L2 acquisition: Off-line and on-line performance. *Bilingualism: Language and Cognition*, *12*(4), 463–483. https://doi.org/10.1017/S1366728909990253

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, *120*(4), 901–931. https://doi.org/10.1016/j.lingua.2009.06.004

Hopp, H. (2017). Cross-linguistic lexical and syntactic co-activation in L2 sentence processing. *Linguistic Approaches to Bilingualism*, *7*(1), 96–130. https://doi.org/10.1075/lab.14027.hop

Houwer, A. de. (1990). *The acquisition of two languages from birth*. Cambridge University Press. https://doi.org/10.1017/CBO9780511519789

Hoyt, W. T., & Del Re, A. C. (2018). Effect size calculation in meta-analyses of psychotherapy outcome research. *Psychotherapy Research*, *28*(3), 379–388. https://doi.org/10.1080/10503307.2017.1405171

Hsin, L., Legendre, G., & Omaki, A. (2013). Priming cross-linguistic interference in Spanish-English bilingual children. In S. Baiz, N. Goldman, & R. Hawkes (Eds.), *Proceedings of the 37th annual Boston University on Language Development (BUCLD 37, Vol. 1)* (pp. 165–177). Cascadilla Press.

Hulk, A., & Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition*, *3*(3), 227–244. https://doi.org/10.1017/S1366728900000353

Järvikivi, J., Pyykkönen-Klauck, P., Schimke, S., Colonna, S., & Hemforth, B. (2014). Information structure cues for 4-year-olds and adults: tracking eye movements to visually presented anaphoric referents. *Language, Cognition and Neuroscience*, *29*(7), 877–892.

https://doi.org/10.1080/01690965.2013.804941

Järvikivi, J., van Gompel, R. P. G., Hyönä, J., & Bertram, R. (2005). Ambiguous pronoun resolution: Contrasting the first-mention and subject-preference accounts. *Psychological Science*, *16*(4), 260–264. https://doi.org/10.1111/j.0956-7976.2005.01525.x

Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning*, *50*(2), 245–309. https://doi.org/10.1111/0023-8333.00118

Jurafsky, D., Bell, A., & Girand, C. (2002). The role of the lemma in form variation. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 3–34). De Gruyter Mouton. https://doi.org/10.1515/9783110197105.1.3

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149. https://doi.org/10.1037/0033-295X.99.1.122

Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin*, *109*(3), 490–501. https://doi.org/10.1037/0033-2909.109.3.490

Kaiser, E. (2011). Salience and contrast effects in reference resolution: The interpretation of Dutch pronouns and demonstratives. *Language and Cognitive Processes*, *26*(10), 1587–1624. https://doi.org/10.1080/01690965.2010.522915

Kaltsa, M., Tsimpli, I. M., & Argyri, F. (2019). The development of gender assignment and agreement in English-Greek and German-Greek bilingual children. *Linguistic Approaches to Bilingualism*, *9*(2), 253–288. https://doi.org/10.1075/lab.16033.kal

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, *14*(2), 185–209. https://doi.org/10.1016/0010-0277(83)90029-X

Kidd, E., Chan, A., & Chiu, J. (2015). Cross-linguistic influence in simultaneous Cantonese–English bilingual children's comprehension of relative clauses. *Bilingualism: Language and Cognition*, *18*(3), 438–452. https://doi.org/10.1017/S1366728914000649

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*(5), 580–602. https://doi.org/10.1016/0749-596X(91)90027-H

Kootstra, G. J., & Doedens, W. J. (2016). How multiple sources of experience influence bilingual syntactic choice: Immediate and cumulative cross-language effects of structural priming, verb bias, and language dominance. *Bilingualism: Language and Cognition*, *19*(4), 710–732. https://doi.org/10.1017/S1366728916000420

Koster, J. (1974). Het werkwoord als spiegelcentrum. *Spektator*, *3*, 601–618.

Koster, J. (1975). Dutch as an SOV Language. *Linguistic Analysis*, *1*, 111–136.

Koutamanis, E., Kootstra, G. J., Dijkstra, T., & Unsworth, S. (2021). Lexical priming as evidence for language-nonselective access in the simultaneous bilingual child's lexicon. In D. Dionne and L.-A. Vidal Covas (Eds.), Proceedings of the 45th annual Boston University Conference on Language Development (*BUCLD 45*) (pp. 413-430). Cascadilla Press.

Kroll, J. F., & Dussias, P. E. (2012). The comprehension of words and sentences in two languages. In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism and multilingualism* (2nd ed., pp. 216–243). John Wiley & Sons. https://doi.org/10.1002/9781118332382.ch9

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), 149–174. https://doi.org/10.1006/jmla.1994.1008

Kupisch, T. (2012). Specific and generic subjects in the Italian of German–Italian simultaneous bilinguals and L2 learners. *Bilingualism: Language and Cognition*, *15*(4), 736–756. https://doi.org/10.1017/S1366728911000691

Kupisch, T. (2014). Adjective placement in simultaneous bilinguals (German–Italian) and the concept of cross-linguistic overcorrection. *Bilingualism: Language and Cognition*, *17*(1), 222–233. https://doi.org/10.1017/S1366728913000382

Kupisch, T., & Barton, D. (2013). Generic reference in adult German bilinguals: How bilinguals deal with variation. *Studia Linguistica*, *67*(1), 1–27. https://doi.org/10.1111/stul.12006

Kupisch, T., Lein, T., Barton, D., Schröder, D. J., Stangen, I., & Stoehr, A. (2014).

Acquisition outcomes across domains in adult simultaneous bilinguals with French as weaker and stronger language. *Journal of French Language Studies*, *24*(3), 347–376. https://doi.org/10.1017/S0959269513000197

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 1–12. https://doi.org/10.3389/fpsyg.2013.00863

Lehmann, W. P. (1978). English: A characteristic SVO language. In W. P. Lehmann (Ed.), *Syntactic typology: Studies in the phenomenology of language* (pp. 169–222). University of Texas Press.

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. https://doi.org/10.3758/s13428-011-0146-0

Lemmerth, N., & Hopp, H. (2019). Gender processing in simultaneous and successive bilingual children: Cross-linguistic lexical and syntactic influences. *Language Acquisition*, *26*(1), 21–45. https://doi.org/10.1080/10489223.2017.1391815

Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. MIT Press.

Levelt, Willem J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–75. https://doi.org/10.1017/S0140525X99001776

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, *17*(3), 673–680. https://doi.org/10.1017/S1366728913000606

Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *The Mental Lexicon*, *3*(3), 349–374. https://doi.org/10.1075/ml.3.3.06lin

Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics*, *41*(5), 791–824. https://doi.org/10.1515/ling.2003.026

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*(5), 605–621. https://doi.org/10.1080/20445911.2013.795574

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)

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97–115. https://doi.org/10.1017/s1366728903001068

Marinis, T., & Armon-Lotem, S. (2015). Sentence repetition. In S. Armon-Lotem, N. Meir, & J. de Jong (Eds.), *Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment* (pp. 116–143). Multilingual Matters.

Marinis, T. (2007). On-line processing of passives in L1 and L2 children. In A. Belikova, L. Meroni, & M. Umeda (Eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)* (pp. 265–276). Cascadilla Proceedings Project.

Marinis, T. (2010). Using on-line processing methods in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 139–162). John Benjamins. https://doi.org/10.1075/lllt.27.09mar

Marinis, T., & Chondrogianni, V. (2011). Comprehension of reflexives and pronouns in sequential bilingual children: Do they pattern similarly to L1 children, L2 adults, or children with specific language impairment? *Journal of Neurolinguistics*, *24*(2), 202–212. https://doi.org/10.1016/j.jneuroling.2010.02.009

Marinis, T., & Saddy, D. (2013). Parsing the passive: Comparing children with specific language impairment to sequential bilingual children. *Language Acquisition*, *20*(2), 155–179. https://doi.org/10.1080/10489223.2013.766743

Martohardjono, G., Phillips, I., Madsen II, C. N., & Schwartz, R. G. (2017).

Cross-linguistic influence in bilingual processing: An ERP study. In M. LaMendola & J. Scott (Eds.), *Proceedings of the 41st annual Boston University Conference on Language Development* (*BUCLD 41*) (pp. 452–465). Cascadilla Press.

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

McLaughlin, B. (1978). *Second language acquisition in childhood*. Erlbaum.

Meir, N., Walters, J., & Armon-Lotem, S. (2017). Bi-directional cross-linguistic influence in bilingual Russian-Hebrew children. *Linguistic Approaches to Bilingualism*, *7*(5), 514–553. https://doi.org/10.1075/lab.15007.mei

Meisel, Jürgen M. (1989). Early differentiation of languages in bilingual children. In K. Hyltenstam & L. Obler (Eds.), *Bilingualism across the lifespan: Aspects of acquisition, maturity and loss* (pp. 13–40). Cambridge University Press. https://doi.org/10.1017/CBO9780511611780.003

Meisel, Jürgen M. (2007). On autonomous syntactic development in multiple first language acquisition. In H. Caunt-Nulton, S. Kulatilake, & I.-H. Woo (Eds.), *Proceedings of the 31th annual Boston University Conference on Language Development (BUCLD 31)* (pp. 26–45). Cascadilla Press.

Meisel, Jürgen M. (2011). *First and second language acquisition: Parallels and differences*. Cambridge University Press. https://doi.org/10.1017/CBO9780511862694

Meroni, L., Smeets, L., & Unsworth, S. (2017). Cross-linguistic influence in scope ambiguity: Evidence for acceleration. In E. Blom, L. Cornips, & J. Schaeffer (Eds.), *Cross-linguistic influence in Bilingualism: In honor of Aafke Hulk* (pp. 181–206). John Benjamins. https://doi.org/10.1075/sibil.52.09mer

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*(1), 25–40. https://doi.org/10.1006/jmla.1998.2602

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), 1–6. https://doi.org/10.1371/journal.pmed.1000097

Montrul, S. (2010). Dominant language transfer in adult second language learners and heritage speakers. *Second Language Research*, *26*(3), 293–327. https://doi.org/10.1177/0267658310365768

Montrul, S. (2011). Multiple interfaces and incomplete acquisition. *Lingua*, *121*(4), 591–604. https://doi.org/10.1016/j.lingua.2010.05.006

Montrul, S., & Ionin, T. (2010). Transfer effects in the interpretation of definite articles by Spanish heritage speakers. *Bilingualism: Language and Cognition*, *13*(4), 449–473. https://doi.org/10.1017/S1366728910000040

Müller, N., & Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition*, *4*(1), 1–21. https://doi.org/10.1017/s1366728901000116

Mykhaylyk, R., & Ytterstad, E. (2017). Directionality of cross-linguistic influence: Which referring choices do bilingual Ukrainian–English children make? *International Journal of Bilingualism*, *21*(1), 99–121. https://doi.org/https://doi.org/10.1177/1367006915603824

Nicoladis, E. (2002). What's the difference between 'toilet paper' and 'paper toilet'? French-English bilingual children's crosslinguistic transfer in compound nouns. *Journal of Child Language*, *29*(4), 843–863. https://doi.org/10.1017/S0305000902005366

Nicoladis, E. (2003). Cross-linguistic transfer in deverbal compounds of preschool bilingual children. *Bilingualism: Language and Cognition*, *6*(1), 17–31. https://doi.org/10.1017/S1366728903001019

Nicoladis, E. (2006). Cross-linguistic transfer in adjective-noun strings by preschool bilingual children. *Bilingualism: Language and Cognition*, *9*(1), 15–32. https://doi.org/10.1017/S136672890500235X

Nicoladis, E. (2012). Cross-linguistic influence in French-English bilingual children's possessive constructions. *Bilingualism: Language and Cognition*, *15*(2), 320–328. https://doi.org/10.1017/S1366728911000101

Nicoladis, E., & Gavrila, A. (2015). Cross-linguistic influence in Welsh-English bilingual children's adjectival constructions. *Journal of Child Language*, *42*(4), 903–916. https://doi.org/https://doi.org/10.1017/S0305000914000440

Nicoladis, E., Rose, A., & Foursha-Stevenson, C. (2010). Thinking for speaking and cross-linguistic transfer in preschool bilingual children. *International Journal of Bilingual Education and Bilingualism*, *13*(3), 345–370. https://doi.org/10.1080/13670050903243043

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing. *Studies in Second Language Acquisition*, *25*(4), 501–528. https://doi.org/10.1017/S0272263103000214

Papadopoulou, D., Peristeri, E., Plemenou, E., Marinis, T., & Tsimpli, I. (2015). Pronoun ambiguity resolution in Greek: Evidence from monolingual adults and children. *Lingua*, *155*, 98–120. https://doi.org/10.1016/j.lingua.2014.09.006

Paradis, J., Emmerzael, K., & Sorenson Duncan, T. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders*, *43*(6), 474–497. https://doi.org/10.1016/j.jcomdis.2010.01.002

Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition*, *18*(1), 1–25. https://doi.org/10.1017/S0272263100014662

Paradis, J., & Navarro, S. (2003). Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: What is the role of the input? *Journal of Child Language*, *30*(2), 371–393. https://doi.org/10.1017/S0305000903005609

Pérez-Leroux, A. T., Cuza, A., & Thomas, D. (2011). Clitic placement in Spanish–English bilingual children. *Bilingualism: Language and Cognition*, *14*(2), 221–232. https://doi.org/10.1017/S1366728910000234

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of*

Memory and Language, *39*(4), 633–651. https://doi.org/10.1006/jmla.1998.2592

Pirvulescu, M., Pérez-Leroux, A.-T., Roberge, Y., Strik, N., & Thomas, D. (2014). Bilingual effects: Exploring object omission in pronominal languages. *Bilingualism: Language and Cognition*, *17*(3), 495–510. https://doi.org/10.1017/S1366728913000631

Poarch, G. J., & van Hell, J. G. (2012). Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, *111*(3), 419–438. https://doi.org/10.1016/j.jecp.2011.09.008

Poeppel, D., & Wexler, K. (1993). The full competence hypothesis of clause structure in early German. *Language*, *69*(1), 1–33. https://doi.org/10.2307/416414

Polinsky, M., & Kagan, O. (2007). Heritage languages: In the 'wild' and in the classroom. *Language and Linguistics Compass*, *1*(5), 368–395. https://doi.org/10.1111/j.1749-818X.2007.00022.x

Polinsky, M., & Scontras, G. (2020). Understanding heritage languages. *Bilingualism: Language and Cognition*, *23*(1), 4–20. https://doi.org/10.1017/S1366728919000245

Polišenská, K., Chiat, S., & Roy, P. (2015). Sentence repetition: what does the task measure? *International Journal of Language & Communication Disorders*, *50*(1), 106–118. https://doi.org/10.1111/1460-6984.12126

Porretta, V., Kyröläinen, A.-J., van Rij, J., & Järvikivi, J. (2020). *VWPre: Tools for preprocessing visual world data* (R package version 1.2.3). Retrieved from https://cran.r-project.org/web/packages/VWPre

Quené, H. (2010). How to design and analyze language acquisition studies. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 269–284). John Benjamins. https://doi.org/10.1075/lllt.27.15que

R Core Team. (2018). *R: A language and environment for statistical computing* (1.0.153) [Computer software]. Retrieved from https://www.r-project.org

R Core Team. (2020). *R: A language and environment for statistical computing* (version 4.0.3) [Computer software]. Retrieved from https://www.r-project.org/

Radford, A. (2004). *Minimalist Syntax: Exploring the structure of English*. Cambridge University Press.

Rinke, E., & Flores, C. (2014). Morphosyntactic knowledge of clitics by Portuguese heritage bilinguals. *Bilingualism: Language and Cognition*, *17*(4), 681–699. https://doi.org/10.1017/S136672891300076X

Rinker, T., & Gagarina, N. (2017). *Cross-linguistic Lexical Task – German*. http://clts.strony.psych.uw.edu.pl/

Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 influence and general learner effects. *Studies in Second Language Acquisition*, *30*(3), 333–357. https://doi.org/10.1017/S0272263108080480

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1–3), 107–142. https://doi.org/10.1016/0010-0277(92)90041-F

Rodina, Y., Kupisch, T., Meir, N., Mitrofanova, N., Urek, O., & Westergaard, M. (2020). Internal and external factors in heritage language acquisition: Evidence from heritage Russian in Israel, Germany, Norway, Latvia and the United Kingdom. *Frontiers in Education*, *5*, 1–17. https://doi.org/10.3389/feduc.2020.00020

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. John Wiley & Sons. https://doi.org/10.1002/0470870168

Runnqvist, E., Gollan, T. H., Costa, A., & Ferreira, V. S. (2013). A disadvantage in bilingual sentence production modulated by syntactic frequency and similarity across languages. *Cognition*, *129*(2), 256–263. https://doi.org/10.1016/j.cognition.2013.07.008

Schepens, J., van der Slik, F., & van Hout, R. (2013). The effect of linguistic distance across Indo-European mother tongues on learning Dutch as a second language. In L. Borin & A. Saxena (Eds.), *Approaches to Measuring Linguistic Differences* (pp. 199–230). De Gruyter. https://doi.org/https://doi.org/10.1515/9783110305258.199

Schimke, S., de la Fuente, I., Hemforth, B., & Colonna, S. (2018). First language influence on second language offline and online ambiguous pronoun resolution. *Language Learning*, *68*(3), 744–779. https://doi.org/10.1111/lang.12293

Schmitt, B. M., Meyer, A. S., & Levelt, W. J. (1999). Lexical access in the production of pronouns. *Cognition*, *69*(3), 313–335. https://doi.org/10.1016/S0010-0277(98)00073-0

Schmitz, K., Di Venanzio, L., & Scherger, A.-L. (2016). Null and overt subjects in Italian and Spanish heritage speakers in Germany. *Lingua*, *180*, 101–123. https://doi.org/10.1016/j.lingua.2016.04.004

Schmitz, K., Patuto, M., & Müller, N. (2012). The null-subject parameter at the interface between syntax and pragmatics: Evidence from bilingual German–Italian, German–French and Italian–French children. *First Language*, *32*(1–2), 205–238. https://doi.org/10.1177/0142723711403880

Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Psychology Software Tools.

Schneider, W. (2015). *Memory development from early childhood through emerging adulthood*. Springer International. https://doi.org/10.1007/978-3-319-09611-7

Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, *56*(2), 153–171. https://doi.org/10.1016/j.jml.2006.10.002

Schwartz, M., Minkov, M., Dieser, E., Protassova, E., Moin, V., & Polinsky, M. (2015). Acquisition of Russian gender agreement by monolingual and bilingual children. *International Journal of Bilingualism*, *19*(6), 726–752. https://doi.org/10.1177/1367006914544989

Serratrice, L. (2007). Cross-linguistic influence in the interpretation of anaphoric and cataphoric pronouns in English-Italian bilingual children. *Bilingualism: Language and Cognition*, *10*(3), 225–238. https://doi.org/10.1017/S1366728907003045

Serratrice, L. (2013). Cross-linguistic influence in bilingual development: Determinants and mechanisms. *Linguistic Approaches to Bilingualism*, *3*(1), 3–25. https://doi.org/10.1075/lab.3.1.01ser

Serratrice, L. (2016). Cross-linguistic influence, cross-linguistic priming and the nature of shared syntactic structures. *Linguistic Approaches to Bilingualism*, *6*(6), 822–827. https://doi.org/10.1075/lab.6.6.15ser

Serratrice, L., Sorace, A., Filiaci, F., & Baldo, M. (2009). Bilingual children's sensitivity to specificity and genericity: Evidence from metalinguistic awareness. *Bilingualism: Language and Cognition*, *12*(2), 239–257. https://doi.org/10.1017/S1366728909004027

Serratrice, L., Sorace, A., Filiaci, F., & Baldo, M. (2012). Pronominal objects in English-Italian and Spanish-Italian bilingual children. *Applied Psycholinguistics*, *33*(4), 725–751. https://doi.org/10.1017/S0142716411000543

Serratrice, L., Sorace, A., & Paoli, S. (2004). Crosslinguistic influence at the syntax–pragmatics interface: Subjects and objects in English–Italian bilingual and monolingual acquisition. *Bilingualism: Language and Cognition*, *7*(3), 183–205. https://doi.org/10.1017/S1366728904001610

Silva-Corvalán, C., & Treffers-Daller, J. (Eds.) (2015). *Language dominance in bilinguals: Issues of measurement and operationalization*. Cambridge University Press. https://doi.org/doi:10.1017/CBO9781107375345

Song, H.-J., & Fisher, C. (2005). Who's "she"? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language*, *52*(1), 29–57. https://doi.org/10.1016/j.jml.2004.06.012

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1–33. https://doi.org/10.1075/lab.1.1.01sor

Sorace, A., & Serratrice, L. (2009). Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, *13*(2), 195–210. https://doi.org/10.1177/1367006909339810

Sorace, A., Serratrice, L., Filiaci, F., & Baldo, M. (2009). Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, *119*(3), 460–477. https://doi.org/10.1016/j.lingua.2008.09.008

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., … Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, *343*(d4002), 1–8. https://doi.org/10.1136/bmj.d4002

Strik, N., & Pérez-Leroux, A. T. (2011). Jij doe wat girafe?: Wh-movement and inversion in Dutch-French bilingual children. *Linguistic Approaches to Bilingualism*, *1*(2), 175–205. https://doi.org/10.1075/lab.1.2.03str

Syrett, K., Lingwall, A., Perez-Cortes, S., Austin, J., Sánchez, L., Baker, H., … Arias-Amaya, A. (2017). Differences between Spanish monolingual and Spanish-English bilingual children in their calculation of entailment-based scalar implicatures. *Glossa: A Journal of General Linguistics*, *2*(1), 1–19. https://doi.org/10.5334/gjgl.76

Tremblay, A., & Ransijn, J. (2020). *Model selection and post-hoc analysis for (G)LMER models* (R package version 3.0). https://cran.r-project.org/package=LMERConvenienceFunctions

Tuller, L. (2015). Clinical use of parental questionnaires in multilingual contexts. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 301–330). Multilingual Matters. https://doi.org/10.21832/9781783093137-013

Tyler, L. K. (1983). The development of discourse mapping processes: The on-line interpretation of anaphoric expressions. *Cognition*, *13*(3), 309–341. https://doi.org/10.1016/0010-0277(83)90013-6

Ünal, Ö., Tunçer, A. M., & Ege, P. (2012). *Cross-linguistic Lexical Task - Turkish*. http://clts.strony.psych.uw.edu.pl/

Unsworth, S. (2003). Testing Hulk & Müller (2000) on crosslinguistic influence: Root Infinitives in a bilingual German/English child. *Bilingualism: Language and Cognition*, *6*(2), 143–158. https://doi.org/10.1017/S1366728903001093

Unsworth, S. (2005). *Child L2, adult L2, child L1: Differences and similarities. A study on the acquisition of direct object scrambling in Dutch* [Doctoral dissertation, Utrecht University]. LOT.

https://www.lotpublications.nl/child-l2-adult-l2-child-l1-differences-and-similarities-child-l2-adult-l2-child-l1-differences-and-similarities-a-study-on-the-acquisition-of-direct-object-scrambling-in-dutch

Unsworth, S. (2012). Testing for crosslinguistic influence and exposure effects in the bilingual acquisition of specific indefinite objects. In A. K. Biller, E. Y. Chung, & A. E. Kimball (Eds.), *Proceedings of the 36th annual Boston University Conference on Language Development (BUCLD 36)* (pp. 633–645). Cascadilla Press.

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition*, *16*(1), 86–110. https://doi.org/10.1017/S1366728912000284

Unsworth, S. (2016). Early child L2 acquisition: Age or input effects? Neither, or both? *Journal of Child Language*, *43*(3), 608–634. https://doi.org/10.1017/S030500091500080X

Unsworth, S., Brouwer, S., de Bree, E., & Verhagen, J. (2019). Predicting bilingual preschoolers' patterns of language development: Degree of non-native input matters. *Applied Psycholinguistics*, *40*(5), 1189–1219. https://doi.org/10.1017/S0142716419000225

Unsworth, S., Chondrogianni, V., & Skarabela, B. (2018). Experiential measures can be used as a proxy for language dominance in bilingual language acquisition research. *Frontiers in Psychology*, *9*, 1–15. https://doi.org/10.3389/fpsyg.2018.01809

van Dijk, C.N., van Wonderen, Koutamanis, E., Kootstra, G.J., & Unsworth, S. (2021). Effect sizes of experimental studies on cross-linguistic influence in bilingual children. DANS. https://doi.org/10.17026/dans-z6a-yw53

van Koert, M., Koeneman, O., Hulk, A., & Weerman, F. (2016). Revisiting binding in Turkish-Dutch bilingual children: Comprehension and production. In D. Stringer, J. Garrett, B. Halloran, & S. Mossman (Eds.), *Proceedings of the 13th Generative Approaches to Second Language Acquisition Conference (GASLA 2015)* (pp. 80–94). Cascadilla Proceedings Project.

van Rij, J., Hendriks, P., van Rijn, H., Baayen, R. H., & Wood, S. N. (2019). Analyzing the time course of pupillometric data. *Trends in Hearing*, *23*, 1–22. https://doi.org/10.1177/2331216519832483 van Rij, J., Wieling, M., Baayen, R. H., & van Rijn, H. (2020). *itsadug: Interpreting time series and autocorrelated data using GAMMs* (R package version 2.4). https://cran.r-project.org/web/packages/itsadug

van Wonderen, E., Blom, E., Boerma, T., van Dijk, C., Janssen, B., & Unsworth, S. (2017). *Cross-linguistic Lexical Task – Dutch*. http://clts.strony.psych.uw.edu.pl/

Vasilyeva, M., Waterfall, H., Gámez, P. B., Gómez, L. E., Bowers, E., & Shimpi, P. (2010). Cross-linguistic syntactic priming in bilingual children. *Journal of Child Language*, *37*(5), 1047–1064. https://doi.org/10.1017/S0305000909990213

Verhagen, A. (1992). Praxis of linguistics: Passives in Dutch. *Cognitive Linguistics*, *3*(3), 301–342. https://doi.org/10.1515/cogl.1992.3.3.301

Verrips, M. (1996). Potatoes must peel: The acquisition of the Dutch passive [Doctoral thesis]. University of Amsterdam. https://hdl.handle.net/11245/1.119921

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Von Holzen, K., Fennel, C. T., & Mani, N. (2019). The impact of cross-language phonological overlap on bilingual and monolingual toddlers' word recognition. *Bilingualism: Language and Cognition*, *22*(3), 476–499. https://doi.org/10.1017/S1366728918000597

Von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*, *113*(4), 569–586. https://doi.org/10.1016/j.jecp.2012.08.001

Vosse, T., & Kempen, G. (2009). The Unification Space implemented as a localist neural net: Predictions and error-tolerance in a constraint-based parser. *Cognitive Neurodynamics*, *3*(4), 331–346. https://doi.org/10.1007/s11571-009-9094-0

Wechsler, D., & Naglieri, J. A. (2008). *Wechsler Nonverbal Scale of Ability*. Pearson Assessment and Information.

Wijnen, F., & Verrips, M. (1998). The acquisition of Dutch syntax. In S. Gillis & A. De Houwer (Eds.), *The acquisition of Dutch* (pp. 233–299). John Benjamins. https://doi.org/10.1075/pbns.52

Witzel, J., Witzel, N., & Nicol, J. (2012). Deeper than shallow: Evidence for structure-based parsing biases in second-language sentence processing. *Applied Psycholinguistics*, *33*(2), 419–456. https://doi.org/10.1017/S0142716411000427

Wood, S. N. (2017). *Generalized additive models: An introduction with R* (2nd ed.). Chapman and Hall/CRC.

Wood, S. N. (2020). *mgcv: Mixed GAM computation vehicle with automatic smoothness estimation* (R package version 1.8-33). https://cran.r-project.org/web/packages/mgcv

Yip, V., & Matthews, S. (2000). Syntactic transfer in a Cantonese-English bilingual child. *Bilingualism: Language and Cognition*, *3*(3), 193–208. https://doi.org/10.1017/S136672890000033X

Yip, V., & Matthews, S. (2006). Assessing language dominance in bilingual acquisition: A case for mean length utterance differentials. *Language Assessment Quarterly*, *3*(2), 97–116. https://doi.org/10.1207/s15434311laq0302_2

Zink, I., & Lejaegere, M. (2002). *N-CDIs: Lijsten voor Communicatieve Ontwikkeling*. Acco.

Zwart, J. W. (2011). *The Syntax of Dutch*. Cambridge University Press. https://doi.org/10.1017/CBO9780511977763

Appendix Chapter 2

A2.1. Narrow versus broad scope surface overlap

Studies differ in how narrowly or broadly they have defined surface overlap. Some authors restricted their assessment of surface overlap to specific contexts, whereas if the contextual scope were broadened, a different assessment would have been made. In many cases, however, it was unclear how narrowly or broadly authors defined the scope of surface overlap. This meant that predictions about surface overlap were not consistent across studies, and in some cases a datapoint that was considered a situation of *no* surface overlap in one study, could have been considered a situation of partial surface overlap according to its definition in another study (see Bernardini & van de Weijer, 2017, for a discussion of their predictions based on both a narrow and broad scope).

We will illustrate the difference in defining surface overlap and its consequences by a study that used a more narrow scope (Argyri & Sorace, 2007) and a study that used a more broad scope (Bosch & Unsworth, 2020).

Argyri & Sorace (2007)

Argyri and Sorace (2007) have studied cross-linguistic influence in various morpho-syntactic properties in both Greek and English. One of these properties is the placement of subjects in *what*-embedded interrogatives in Greek. In Greek the subject in *what*-embedded interrogatives can only be placed in postverbal position (1, example taken from Argyri & Sorace, 2007; p. 89) and not in preverbal position (2). In contrast, in English the subject in *what*-embedded interrogatives can only be placed in preverbal position (3) and not in postverbal position (4). These differences in word order between Greek and English in the context of *what*-embedded interrogatives made the authors define the situation in Greek as a situation of *no* surface overlap. Hence, surface overlap in this study was narrowly defined.

(1) I jaja den thimate ti efage <u>o Nikos</u>.
the grandmother not remember what ate the Nikos
"Grandmother doesn't remember what Nikos ate."

(2) *I jaja den thimate ti <u>o Nikos</u> efage.
the grandmother not remember what the Nikos ate
"Grandmother doesn't remember what Nikos ate."

(3) The grandmother does not remember what **Nikos** ate.

(4) *The grandmother does not remember what ate **Nikos**.

However, surface overlap can also be more broadly defined in Argyri and Sorace's study. This results in a different categorization of overlap. Although preverbal subjects are ungrammatical in *what*-embedded interrogatives, preverbal subjects are allowed in other sentence contexts in Greek (as illustrated by the preverbal position of "the grandmother" in 1). Hence, in a broader sense Greek allows for both preverbal and postverbal subjects. Therefore, it could be argued that with respect to subject placement in *what*-embedded interrogatives in Greek there is a situation of *surface* overlap with English: Greek provides evidence for two options – preverbal and postverbal subjects – and English might reinforce the incorrect option – preverbal subject placement.

Bosch & Unsworth (2020)

An example of a study with a broader interpretation of surface overlap is Bosch and Unsworth (2020). They studied English-Dutch children's production and judgements of verb placement in English. Dutch is a verb-second language meaning that in main clauses the verb occurs in second position (5 and 6). In contrast, English has SVO order, meaning that the subject should precede the verb (7). A sentence structure like in (6) would be ungrammatical. An exception to this rule are *wh*-questions. In these questions, the auxiliary can appear in second position in front of the verb (8).

(5) Zij **tekende** een auto.
she painted a car
"She painted a car."

(6) Gisteren **tekende** zij een auto.
yesterday painted she a car
"Yesterday she painted a car."

(7) Yesterday she **painted** a car.

(8) Where **are** you?

Bosch and Unsworth asked children to produce and judge main clauses with a main verb only (like in 7) or with an auxiliary and a main verb (like in 9). In both type of clauses, the main verb and auxiliary should appear after the subject in third position. Hence, in this specific context there is no surface overlap between English (verb third) and Dutch (verb second). However, the authors apply a broader definition of surface overlap. They argue that there is optionality in English in structures with an auxiliary, because auxiliaries sometimes appear in preverbal position in *wh*-questions. Therefore, the authors consider the structures with an auxiliary in English a situation of *surface* overlap and they predict that Dutch might reinforce the incorrect verb second structure in this context.

(9) Today she **is** painting a car.

Crucially, in the studies by Argyri and Sorace (2007) and Bosch and Unsworth (2020) evidence of cross-linguistic influence in Greek *what*-embedded interrogatives and in English main clauses with an auxiliary can be interpreted as evidence *for* or *against* the surface overlap hypothesis. This all depends on how narrow or broad surface overlap is defined. If the authors' definitions are applied, cross-linguistic influence in Greek *what*-embedded interrogatives would offer support against the surface overlap hypothesis, whereas cross-linguistic influence in English main clauses with an auxiliary would offer support in favour of the surface overlap hypothesis. However, the situation would be reversed if Argyri and Sorace's narrow definition of surface overlap would be applied on Bosch and Unsworth's study or if Bosch and Unsworth's broad definition of surface overlap would be applied on Argyri and Sorace's study. This shows that how surface overlap is defined within a study can have a crucial impact on the interpretation of evidence of cross-linguistic influence.

A2.2. Explanation coding procedure surface overlap based on the adult language system

In this section we give a detailed illustration of how we classified datapoints as *partial overlap*, *no overlap* or *complete overlap* based on Nicoladis (2006). In this study, Nicoladis tested for French-English children's placement of adjectives. In French, adjectives can be in pre- and postnominal position. However, certain adjectives are typically prenominal (e.g., *le grand*$_{ADJ}$ *chat*$_{N}$/the big cat) and others are typically postnominal (e.g., *le chat*$_{ADJ}$ *noir*$_{N}$/the cat black). In English, on the other hand, adjectives are prenominal (e.g., *the big/black cat*). Nicoladis elicited adjective-noun strings in both French and English. She reported results separately for adjectives that were typically placed in prenominal and postnominal position in French. In other words, she reported on bilingual and monolingual children's behaviour in four situations: the placement of adjectives that typically occur in prenominal position in French, both in French production data and in English production data, and the placement of adjectives that typically occur in postnominal position in French, again both in French production data and in English production data (see Tables A2.2.1 and A2.2.2).

Table A2.2.1. Situation of adjectives that typically occur in postnominal position in French

				Overlap situation	
	French		**English**	**English-to-French**	**French-to-English**
Target	Postnominal *le chat noir*	→	*Postnominal **the cat black*	Partial overlap	No overlap
Non-target	Prenominal *?le noir chat*	←	Prenominal *the black cat*		

Table A2.2.2. Situation of adjectives that typically occur in prenominal position in French

				Overlap situation	
	French		**English**	**English-to-French**	**French-to-English**
Target	Prenominal *le grand chat*	← →	Prenominal *the big cat*	Partial overlap	Complete overlap
Non-target	Postnominal *?le chat grand*				

A datapoint was coded as a situation of *no overlap* when children's languages allowed for only one option and these options differed between languages. In this situation, we predicted that, if cross-linguistic influence was present, it should result in a more frequent choice of the option from the non-target language in bilingual children in the language tested than their monolingual peers. In Nicoladis (2006), a situation of *no overlap* occurred in children's placement of adjectives in English that were typically in postnominal position in French: in English, adjectives should always be in prenominal position, whereas in French, these adjectives should be in postnominal position (Table A2.2.1). We therefore predicted that cross-linguistic influence could result in a stronger use of postnominal adjectives in English by bilingual French-English children than by monolingual English children.

A datapoint was coded as a situation of *partial overlap* in two situations: (i) when the language tested allowed for two options and children's other language allowed for only one of these options, and (ii) when both bilingual children's languages allowed for the same two options, but differed in their preferences for one of these options. We predicted the preferred or only option in one language to reinforce the overlapping option in the other language in case of cross-linguistic influence. In Nicoladis (2006), this concerns the placement of prenominal and postnominal adjectives in French (see Table A2.2.1 and Table A2.2.2). Hence, we predicted cross-linguistic influence from English to French, reinforcing the correct placement of French prenominal adjectives in prenominal position and the incorrect placement of French postnominal adjectives in prenominal position.

Finally, a datapoint was coded as a situation of *complete overlap* if morphosyntactic properties were similar between children's languages in the specific context tested. In these situations we predicted no cross-linguistic influence. With respect to Nicoladis (2006), this concerns the placement of adjectives in English that are typically prenominal in French (see Table A2.2.2). Both in English and in French, these adjectives should be placed in prenominal position. Therefore, we coded this situation as *complete overlap* from the perspective of English. Because English only allows for prenominal adjectives, we did not assume that French would further reinforce correct adjective placement (assuming monolingual children would already be performing at ceiling). As a consequence, no cross-linguistic influence was predicted.

A2.3. Descriptives of the studies in the dataset

Table A2.3.1. Characteristics of the studies in the dataset.

Study characteristic	Number of studies	Number of tasks	Number of observations
Task type			
Elicited production	15	23	64
Judgements	10	18	104
Comprehension	4	5	11
Other	2	3	8
Language tested			
English	15	18	72
Italian	6	6	39
Dutch	4	6	18
Greek	1	2	16
French	7	6	14
Spanish	3	3	9
Ukrainian	1	1	6
European Portuguese	1	1	5
Russian	1	1	4
Hebrew	1	1	2
Persian	1	1	2
Language property studied			
Adjective order	3	4	7
Adverb order	1	1	3
Clitic order	3	4	10
Compound order	3	7	9
Object pronoun order	2	4	12
Possessor order	1	2	4
Subject order	1	4	16
Verb order	2	3	10
Wh order	1	1	4
Genericity/specificity	2	4	40
Modified noun	1	2	2
Morpheme marker	2	4	12
Null/overt object pronoun	3	4	14
Null/overt subject pronoun	3	7	28
Pseudo stripping	1	1	2
Quantifiers	1	1	3
Scrambled indefinites	1	3	3
That-trace	1	2	8

Table A2.3.1 (continued).

Study characteristic	Number of studies	Number of tasks	Number of observations
Number of items			
< 5	5	8	50
5-10	13	24	90
10-15	7	11	13
15-20	4	5	13
20-25	5	6	17
>25	1	2	4

Note. The column "Number of studies/tasks" shows how many studies/tasks tested for a certain characteristic. The column "Number of observations" shows how many datapoints in our dataset tested for a certain characteristic. Numbers are based on uncollapsed observations.

Table A2.3.2. Characteristics of the bilingual groups in the dataset.

Child characteristics	Number of studies	Number of groups	Number of observations
Mean age bilingual children			
3-years old	2	2	4
4-years old	12	13	29
5-years old	6	8	23
6-years old	6	5	38
7-years old	3	4	14
8-years old	5	6	68
9-years old	2	3	11
Number of bilingual children			
0-10	6	10	34
10-20	11	13	87
20-30	11	9	38
30-40	7	7	17
> 40	2	2	11

Note. The column "Number of studies" shows how many studies tested groups of children with a certain characteristic. The column "Number of groups" shows how many groups of children with a certain characteristic were tested. The column "Number of observations" shows how many datapoints in our dataset were based on a group of children with a certain characteristic.

Appendix Chapter 3

A3.1. Experimental items with long passive structure used in self-paced listening task (Chapters 3 and 4).

Item	Word order	Segment 1	2	3	4	5	6	7
1	PP-V	de leeuw	wordt	door de beer	gekieteld	en	de muis	rent door het bos
		the lion	*is being*	*by the bear*	*tickled*	*and*	*the mouse*	*runs through the forest*
	V-PP	de leeuw	wordt	gekieteld	door de beer	en	de muis	rent door het bos
		the lion	*is being*	*tickled*	*by the bear*	*and*	*the mouse*	*runs through the forest*
	Question	-						
2	PP-V	de olifant	wordt	door de krokodil	gevoerd	en	de haan	klimt op het dak
		the elephant	*is being*	*by the crocodile*	*fed*	*and*	*the rooster*	*climbes on the roof*
	V-PP	de olifant	wordt	gevoerd	door de krokodil	en	de haan	klimt op het dak
		the elephant	*is being*	*fed*	*by the crocodile*	*and*	*the rooster*	*climbes on the roof*
	Question	-						
3	PP-V	de tijger	wordt	door de ezel	geduwd	en	de aap	springt in de boom
		the tiger	*is being*	*by the donkey*	*pushed*	*and*	*the monkey*	*jumps in the tree*
	V-PP	de tijger	wordt	geduwd	door de ezel	en	de aap	springt in de boom
		the tiger	*is being*	*pushed*	*by the donkey*	*and*	*the monkey*	*jumps in the tree*
	Question	-						
4	PP-V	de wolf	wordt	door de vos	geaaid	en	de pony	schopt tegen de deur
		the wolf	*is being*	*by the fox*	*petted*	*and*	*the pony*	*kicks against the door*
	V-PP	de wolf	wordt	geaaid	door de vos	en	de pony	schopt tegen de deur
		the wolf	*is being*	*petted*	*by the fox*	*and*	*the pony*	*kicks against the door*
	Question	-						

A3.1 (continued).

		Segment						
Item	**Word order**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
5	PP-V	de schilpad	wordt	door de pinguïn	gewassen	en	de papegaai	vliegt door de lucht
		the turtle	*is being*	*by the penguin*	*washed*	*and*	*the parrot*	*flies through the sky*
	V-PP	de schildpad	wordt	gewassen	door de pinguïn	en	de papegaai	vliegt door de lucht
		the turtle	*is being*	*washed*	*by the penguin*	*and*	*the parrot*	*flies through the sky*
	Question	-						
6	PP-V	de koe	wordt	door de geit	vies gemaakt	en	de krokodil	loopt naar de winkel
		the cow	*is being*	*by the goat*	*made dirty*	*and*	*the crocodile*	*walks to the store*
	V-PP	de koe	wordt	vies gemaakt	door de geit	en	de krokodil	loopt naar de winkel
		the cow	*is being*	*made dirty*	*by the goat*	*and*	*the crocodile*	*walks to the store*
	Question	Loopt de krokodil?						
		Is the crocodile walking?						
7	PP-V	de eekhoorn	wordt	door de hamster	gedragen	en	de uil	wandelt door het park
		the squirrel	*is being*	*by the hamster*	*carried*	*and*	*the owl*	*walks through the park*
	V-PP	de eekhoorn	wordt	gedragen	door de hamster	en	de uil	wandelt door het park
		the squirrel	*is being*	*carried*	*by the hamster*	*and*	*the owl*	*walks through the park*
	Question	Wordt de uil gedragen?						
		Is the owl being carried?						
8	PP-V	de egel	wordt	door de kikker	geslagen	en	de geit	tekent een banaan
		the hedgehog	*is being*	*by the frog*	*hit*	*and*	*the goat*	*draws a banana*
	V-PP	de egel	wordt	geslagen	door de kikker	en	de geit	tekent een banaan
		the hedgehog	*is being*	*hit*	*by the frog*	*and*	*the goat*	*draws a banana*
	Question	-						

A3.1 (continued).

Item	Word order	Segment 1	2	3	4	5	6	7
9	PP-V	de giraf	wordt	door de zebra	gebeten	en	de vos	gooit een bal
		the giraffe	*is being*	*by the zebra*	*bitten*	*and*	*the fox*	*throws a ball*
	V-PP	de giraf	wordt	gebeten	door de zebra	en	de vos	gooit een bal
		the giraffe	*is being*	*bitten*	*by the zebra*	*and*	*the fox*	*throws a ball*
	Question	Gooit de vos een bal?						
		Is the fox throwing a ball?						
10	PP-V	de kip	wordt	door de eend	geborsteld	en	de slang	speelt een spel
		the chicken	*is being*	*by the duck*	*brushed*	*and*	*the snake*	*plays a game*
	V-PP	de kip	wordt	geborsteld	door de eend	en	de slang	speelt een spel
		the chicken	*is being*	*brushed*	*by the duck*	*and*	*the snake*	*plays a game*
	Question	Speelt de kip een spel?						
		Is the chicken playing a game?						
11	PP-V	de ezel	wordt	door de giraf	gekrabd	en	de pinguïn	ruikt aan een bloem
		the donkey	*is being*	*by the giraffe*	*scratched*	*and*	*the penguin*	*smells a flower*
	V-PP	de ezel	wordt	gekrabd	door de giraf	en	de pinguïn	ruikt aan een bloem
		the donkey	*is being*	*scratched*	*by the giraffe*	*and*	*the penguin*	*smells a flower*
	Question	-						
12	PP-V	de zebra	wordt	door de pony	geknuffeld	en	de vogel	kijkt naar een stoel
		the zebra	*is being*	*by the pony*	*hugged*	*and*	*the bird*	*looks at a chair*
	V-PP	de zebra	wordt	geknuffeld	door de pony	en	de vogel	kijkt naar een stoel
		the zebra	*is being*	*hugged*	*by the pony*	*and*	*the bird*	*looks at a chair*
	Question	-						

A3.1 (continued).

Item	Word order	Segment 1	2	3	4	5	6	7
13	PP-V	de eend	wordt	door de uil	achtervolgd	en	de schildpad	likt aan een ijsje
		the duck	*is being*	*by the owl*	*followed*	*and*	*the turtle*	*licks an ice cream*
	V-PP	de eend	wordt	achtervolgd	door de uil	en	de schildpad	likt aan een ijsje
		the duck	*is being*	*followed*	*by the owl*	*and*	*the turtle*	*licks an ice cream*
	Question	Likt de schildpad?						
		Is the turtle licking?						
14	PP-V	de beer	wordt	door de wolf	gekamd	en	de cavia	vindt een koekje
		the bear	*is being*	*by the wolf*	*combed*	*and*	*the guinea pig*	*finds a cookie*
	V-PP	de beer	wordt	gekamd	door de wolf	en	de cavia	vindt een koekje
		the bear	*is being*	*combed*	*by the wolf*	*and*	*the guinea pig*	*finds a cookie*
	Question	-						
15	PP-V	de kikker	wordt	door de eekhoorn	gekust	en	de vis	zwemt in het water
		the frog	*is being*	*by the squirrel*	*kissed*	*and*	*the fish*	*swims in the water*
	V-PP	de kikker	wordt	gekust	door de eekhoorn	en	de vis	zwemt in het water
		the frog	*is being*	*kissed*	*by the squirrel*	*and*	*the fish*	*swims in the water*
	Question	-						

A3.2. Experimental items with verb second and verb third structure used in self-paced listening task (Chapters 3 and 4).

		Segment				
Item	Word order	1	2	3	4	5
1	V2	op de handdoek	drinkt	de muis	een slok	van de limonade
		on the towel	*drinks*	*the mouse*	*a sip*	*of the lemonade*
	V3	*op de handdoek	de muis	drinkt	een slok	van de limonade
		on the towel	*the mouse*	*drinks*	*a sip*	*of the lemonade*
	Question	-				
2	V2	op de glijbaan	eet	de haan	een snoepje	uit de zak
		on the slide	*eats*	*the rooster*	*(a) candy*	*out of the bag*
	V3	*op de glijbaan	de haan	eet	een snoepje	uit de zak
		on the slide	*the rooster*	*eats*	*(a) candy*	*out of the bag*
	Question	-				
3	V2	bij het water	bouwt	de aap	een toren	van blokken
		at the water	*builds*	*the monkey*	*a tower*	*of cubes*
	V3	*bij het water	de aap	bouwt	een toren	van blokken
		at the water	*the monkey*	*builds*	*a tower*	*of cubes*
	Question	-				
4	V2	op de stoel	leest	de pony	een verhaaltje	in het boek
		on the chair	*reads*	*the pony*	*a story*	*in the book*
	V3	*op de stoel	de pony	leest	een verhaaltje	in het boek
		on the chair	*the pony*	*reads*	*a story*	*in the book*
	Question	-				

A3.2 (continued).

Item	Word order	Segment 1	2	3	4	5
5	V2	onder de boom	maakt	de papegaai	een foto	van het huis
		underneath the tree	*makes*	*the parrot*	*a picture*	*of the house*
	V3	*onder de boom	maakt	de papegaai	een foto	van het huis
		underneath the tree	*makes*	*the parrot*	*a picture*	*of the house*
	Question	Zit de papegaai in de boom?				
		Is the parrot in the tree?				
6	V2	in de hoek	pakt	de krokodil	een puzzel	uit de kast
		in the corner	*gets*	*the crocodile*	*a puzzle*	*from the shelves*
	V3	*in de hoek	de krokodil	pakt	een puzzel	uit de kast
		in the corner	*the crocodile*	*gets*	*a puzzle*	*from the shelves*
	Question	*Pakt de krokodil iets onder de kast?*				
		Does the crocodile get something from underneath the shelves?				
7	V2	op de steen	poetst	de uil	een knoop	van de jas
		at the rock	*wipes*	*the owl*	*a button*	*of the coat*
	V3	*op de steen	de uil	poetst	een knoop	van de jas
		at the rock	*the owl*	*wipes*	*a button*	*of the coati*
	Question	-				
8	V2	in de oven	bakt	de geit	een taart	van chocola
		in the oven	*bakes*	*the goat*	*a cake*	*of chocolate*
	V3	*in de oven	de geit	bakt	een taart	van chocola
		in the oven	*the goat*	*bakes*	*a cake*	*of chocolate*
	Question	-				

A3.2 (continued).

		Segment				
Item	**Word order**	**1**	**2**	**3**	**4**	**5**
9	V2	op het gras	vouwt	de vos	een hoedje	van papier
		on the grass	*folds*	*the fox*	*a little hat*	*out of paper*
	V3	*op het gras	de vos	vouwt	een hoedje	van papier
		on the grass	*the fox*	*folds*	*a little hat*	*out of paper*
	Question	Vouwt de vos een hoedje van gras?				
		Does the fox fold a little hat out of grass?				
10	V2	achter het hek	zoekt	de hamster	een ei	van de kip
		behind the gate	*searches*	*the hamster*	*an egg*	*from the chicken*
	V3	*achter het hek	de hamster	zoekt	een ei	van de kip
		behind the gate	*the hamster*	*searches*	*an egge*	*from the chicken*
	Question	-				
11	V2	onder de kast	verstopt	de pinguïn	een blokje	van de lego
		underneath the closet	*hides*	*the penguin*	*a cube*	*of (the) lego*
	V3	*onder de kast	de pinguïn	verstopt	een blokje	van de lego
		underneath the closet	*the penguin*	*hides*	*a cube*	*of (the) lego*
	Question	Verstopt de pinguïn iets onder de kast?				
		Does the penguin hide something underneath the closet?				
12	V2	op de bank	zingt	de slang	een liedje	van school
		at the couch	*sings*	*the snake*	*a song*	*from school*
	V3	*op de bang	de slang	zingt	een liedje	van school
		at the couch	*the snake*	*sings*	*a song*	*from school*
	Question	-				

A3.2 (continued).

		Segment				
Item	Word order	1	2	3	4	5
13	V2	onder de tafel	veegt	de vogel	een stukje	van de vloer
		below the table	*brushes*	*the bird*	*a piece*	*of the floor*
	V3	*onder de tafel	de vogel	veegt	een stukje	van de vloer
		below the table	*the bird*	*brushes*	*a piece*	*of the floor*
	Question	Veegt de vogel onder de tafel?				
		Does the bird brush below the table?				
14	V2	op de schommel	knoeit	de cavia	een boterham	van het bordje
		at the swing	*spills*	*the guinea pig*	*a sandwich*	*from the little plate*
	V3	*op de schommel	de cavia	knoeit	een boterham	van het bordje
		at the swing	*the guinea pig*	*spills*	*a sandwich*	*from the little plate*
	Question	-				
15	V2	op de grond	schrijft	de vis	een letter	van zijn naam
		at the floor	*writes*	*the fish*	*a character*	*of his name*
	V3	*op de grond	de vis	schrijft	een letter	van zijn naam
		at the floor	*the fish*	*writes*	*a character*	*of his name*
	Question	-				

A3.3. Children's mean RTs (and SDs) in ms on the self-paced listening task per segment in the PP-V and V-PP condition for each group

		Segment						
		1 **De leeuw** *the lion*	**2** **wordt** *is*	**3** **PP-V:** **door de beer** *by the bear* **V-PP:** **geduwd** *pushed*	**4** **PP-V:** **geduwd** *pushed* **V-PP:** **door de beer** *by the bear*	**5** **en** *and*	**6** **de muis** *the mouse*	**7** **likt** **(…)** *licks*
Dutch monolingual	PP-V	1339 (334)	1108 (301)	1392 (351)	1423 (395)	1081 (295)	1296 (334)	1687 (585)
	V-PP	1355 (336)	1107 (312)	1346 (349)	1429 (359)	1038 (301)	1301 (318)	1700 (512)
English-Dutch	PP-V	1281 (323)	1019 (278)	1313 (329)	1338 (331)	1003 (264)	1252 (331)	1691 (681)
	V-PP	1303 (348)	1017 (255)	1279 (312)	1370 (321)	980 (266)	1237 (297)	1663 (630)
German-Dutch	PP-V	1225 (310)	974 (247)	1281 (309)	1368 (407)	972 (271)	1224 (315)	1649 (676)
	V-PP	1232 (330)	979 (249)	1238 (303)	1343 (346)	980 (257)	1228 (325)	1596 (653)

A3.4. Summary of the mixed linear model of English-Dutch, German-Dutch and monolingual Dutch children's log residual RTs on the long passive sentences. The reference level is German-Dutch children's log residual RTs at segment 4 in the PP-V condition.[a]

Segment*Word order*Group + Duration + RT previous trial + Trial number + Forward Digit Span + Age + (1 + Word order|ID) + (1|Item)

Random effects		*Var*	*SD*
ID	Intercept	0.0011	0.0332
	Word order	<0.0001	0.0037
Item	Intercept	<0.0001	0.0055
Residual		0.0029	0.0539

Fixed effects	*B*	*SE*	*t*	*p*
Intercept	3.2088	0.0061	529.9	<.001
Segment 2	-0.0631	0.0043	-14.8	<.001
Segment 3	-0.0342	0.0032	-10.6	<.001
Segment 5	-0.0532	0.0053	-10.0	<.001
Word order(V-PP)	-0.0092	0.0032	-2.8	.005
Group(English-Dutch)	0.0038	0.0083	0.5	.646
Group(Dutch)	0.0120	0.0082	1.5	.146
Segment 2*Word order(V-PP)	0.0114	0.0045	2.6	.010
Segment 3*Word order(V-PP)	0.0018	0.0046	0.4	.702
Segment 5*Word order(V-PP)	0.0169	0.0044	3.8	<.001
Segment 2*Group(English-Dutch)	0.0083	0.0046	1.8	.071
Segment 3*Group(English-Dutch)	0.0071	0.0047	1.5	.128
Segment 5*Group(English-Dutch)	0.0099	0.0046	2.1	.032
Segment 2*Group(Dutch)	0.0186	0.0046	4.1	<.001
Segment 3*Group(Dutch)	0.0122	0.0046	2.6	.008
Segment 5*Group(Dutch)	0.0138	0.0046	3.0	.003
Word order(V-PP*Group(English-Dutch)	0.0096	0.0047	2.0	.042
Word order(V-PP*Group(Dutch)	0.0105	0.0047	2.2	.025
Segment 2*Word order(V-PP)*Group(English-Dutch)	-0.0110	0.0065	-1.7	.090
Segment 3*Word order(V-PP)*Group(English-Dutch)	-0.0062	0.0066	-0.9	.347
Segment 5*Word order(V-PP)*Group(English-Dutch)	-0.0215	0.0065	-3.3	<.001
Segment 2*Word order(V-PP)*Group(Dutch)	-0.0149	0.0065	-2.3	.022
Segment 3*Word order(V-PP)*Group(Dutch)	-0.0101	0.0065	-1.5	.123
Segment 5*Word order(V-PP)*Group(Dutch)	-0.0270	0.0065	-4.2	<.001
Duration	-0.0002	<0.0001	-27.5	<.001
RT previous trial	0.0001	<0.0001	35.8	<.001

A3.4 (continued).

Fixed effects	B	SE	t	p
Trial number	-0.0006	0.0001	-5.3	<.001
Forward digit span	-0.0012	0.0010	-1.2	.231
Age	-0.0094	0.0026	-3.6	<.001

[a]Note that differences in the estimates between segments do not directly correspond to differences in children's residual listening times between segments. This is because the model controls for the effect of *Duration* on children's residual listening times. That is, the longer the *Duration* of the audio fragment, the shorter the corresponding residual listening time.

A3.5. Summary of the mixed linear model of English-Dutch, German-Dutch and monolingual Dutch children's log residual RTs on the long passive sentences. The reference level is German-Dutch children's log residual RTs at segment 5 in the PP-V condition.[a]

Segment*Word order*Group + Duration + RT previous trial + Trial number + Forward Digit Span + Age + (1 + Word order|ID) + (1|Item)

Random effects		Var	SD
ID	Intercept	0.0011	0.0332
	Word order	< 0.0001	0.0037
Item	Intercept	< 0.0001	0.0055
Residual		0.0029	0.0539

Fixed effects	B	SE	t	p
Intercept	3.1557	0.0063	497.6	<.001
Segment 2	-0.0099	0.0034	-2.9	.004
Segment 3	0.0189	0.0058	3.3	.001
Segment 4	0.0532	0.0053	10.0	<.001
Word order(V-PP)	0.0078	0.0031	2.5	.013
Group(English-Dutch)	0.0137	0.0083	1.6	.102
Group(Dutch)	0.0257	0.0081	3.2	.002
Segment 2*Word order(V-PP)	-0.0055	0.0044	-1.3	.209
Segment 3*Word order(V-PP)	-0.0152	0.0045	-3.4	<.001
Segment 4*Word order(V-PP)	-0.0169	0.0044	-3.8	<.001
Segment 2*Group(English-Dutch)	-0.0015	0.0045	-0.3	.733
Segment 3*Group(English-Dutch)	-0.0028	0.0046	-0.6	.546
Segment 4*Group(English-Dutch)	-0.0099	0.0046	-2.1	.032
Segment 2*Group(Dutch)	0.0048	0.0045	1.1	.288
Segment 3*Group(Dutch)	-0.0015	0.0046	-0.3	.737

A3.5 (continued).

Fixed effects	*B*	*SE*	*t*	*p*
Segment 4*Group(Dutch)	-0.0138	0.0046	-3.0	.003
Word order(V-PP)*Group(English-Dutch)	-0.0118	0.0046	-2.6	.010
Word order(V-PP)*Group(Dutch)	-0.0165	0.0046	-3.6	<.001
Segment 2*Word order(V-PP)*Group(English-Dutch)	0.0104	0.0064	1.6	.103
Segment 3*Word order(V-PP)*Group(English-Dutch)	0.0153	0.0065	2.4	.018
Segment 4*Word order(V-PP)*Group(English-Dutch)	0.0215	0.0065	3.3	<.001
Segment 2*Word order(V-PP)*Group(Dutch)	0.0121	0.0064	1.9	.058
Segment 3*Word order(V-PP)*Group(Dutch)	0.0170	0.0064	2.6	.008
Segment 4*Word order(V-PP)*Group(Dutch)	0.0270	0.0065	4.2	<.001
Duration	-0.0002	<0.0001	-27.5	<.001
RT previous trial	0.0001	<0.0001	35.8	<.001
Trial number	-0.0006	0.0001	-5.3	<.001
Forward digit span	-0.0012	0.0010	-1.2	.231
Age	-0.0094	0.0026	-3.6	<.001

[a]Note that differences in the estimates between segments do not directly correspond to differences in children's residual listening times between segments. This is because the model controls for the effect of *Duration* on children's residual listening times. That is, the longer the *Duration* of the audio fragment, the shorter the corresponding residual listening time.

A3.6. Children's mean RTs (and SDs) in ms on the self-paced listening task per segment in the V2 and V3 condition for each group.

		Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
		op de bank *on the couch*	**V2: zingt** *sings* **V3: de slang** *the snake*	**V2: de slang** *the snake* **V3: zingt** *sings*	**een liedje** a song	**van school** from school
Dutch monolingual	**V2**	1412 (349)	1112 (299)	1254 (315)	1266 (297)	1560 (427)
	V3	1404 (353)	1292 (333)	1126 (296)	1270 (322)	1522 (416)
English-Dutch	**V2**	1340 (357)	1039 (253)	1169 (278)	1195 (274)	1584 (492)
	V3	1332 (341)	1226 (325)	1059 (280)	1210 (314)	1549 (464)
German-Dutch	**V2**	1291 (345)	1015 (257)	1163 (310)	1181 (295)	1573 (514)
	V3	1291 (347)	1169 (303)	1038 (291)	1191 (301)	1563 (525)

A3.7. Summary of the mixed linear model of English-Dutch, German-Dutch and monolingual Dutch children's log residual RTs on the V2/V3 sentences with monolingual Dutch children's log residual RTs at segment 1 in the V2 condition as reference level (Chapter 3).[a]

Segment * Word order + Group + Duration + RT previous trial + Trial number + Forward Digit Span + Age + (1 + Word order|ID) + (1|Item)					
Random effects		*Variance*		*SD*	
ID	Intercept	0.0021		0.0463	
	Word order	0.0001		0.0100	
Item	Intercept	< 0.0001		0.0068	
Residuals		0.0044		0.0662	
Fixed effects		*B*	*SE*	*t*	*p*
Intercept		3.0890	0.0081	379.481	<.001
Segment 2		0.0146	0.0040	3.653	<.001
Segment 3		0.0245	0.0030	8.063	<.001
Segment 4		0.0336	0.0030	11.065	<.001
Word order – V3		0.0018	0.0025	0.701	.483
Group – English-Dutch		-0.0247	0.0109	-2.26	.026
Group – German-Dutch		-0.0359	0.0105	-3.426	<.001
Segment 2 * Word order – V3		0.0212	0.0036	5.825	<.001
Segment 3 * Word order – V3		-0.0142	0.0035	-4.086	<.001
Segment 4 * Word order – V3		-0.0030	0.0033	-0.905	.365
Duration		-0.0002	< 0.0001	-27.000	<.001
RT previous trial		< 0.0001	< 0.0001	30.010	<.001
Trial number		-0.0009	0.0001	-6.951	<.001
Forward digit span		-0.0013	0.0014	-0.888	.376
Age		-0.0139	0.0037	-3.722	<.001

[a]Note that differences in the estimates between segments do not directly correspond to differences in children's residual listening times between segments. This is because the model controls for the effect of *Duration* on children's residual listening times. That is, the longer the *Duration* of the audio fragment, the shorter the corresponding residual listening time.

Appendix Chapter 4

A4.1. Participants' mean residual RTs and SDs on the self-paced listening task by test session by sentence type.

Table A4.1.1. Participants' mean residual RTs (and SDs) in ms on the self-paced listening task per segment in the PP-V and V-PP condition for each group in session 1 (monolingual session).

		Segment						
		1 **De leeuw** *the lion*	2 **wordt** *is*	3 PP-V: **door de beer** *by the bear* V-PP: **geduwd** *pushed*	4 PP-V: **geduwd** *pushed* V-PP: **door de beer** *by the bear*	5 **en** *and*	6 **de muis** *the mouse*	7 **likt** **(…)** *licks*
Dutch monolingual	PP-V	301 (207)	326 (185)	223 (192)	309 (262)	396 (190)	300 (203)	350 (393)
	V-PP	278 (217)	378 (231)	213 (204)	304 (275)	433 (225)	307 (195)	337 (362)
English-Dutch	PP-V	339 (266)	403 (290)	302 (347)	440 (383)	487 (282)	336 (243)	411 (428)
	V-PP	349 (277)	423 (267)	284 (311)	362 (326)	506 (277)	375 (280)	421 (408)
German-Dutch	PP-V	299 (233)	337 (210)	235 (256)	346 (319)	439 (228)	324 (255)	326 (370)
	V-PP	296 (246)	397 (249)	213 (236)	287 (253)	455 (253)	341 (242)	353 (380)

Table A4.1.2. Participants' mean residual RTs (and SDs) in ms on the self-paced listening task per segment in the V2 and V3 condition for each group in session 1 (monolingual session).

		Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
		op de bank *on the couch*	**V2: zingt** *sings* **V3: de slang** *the snake*	**V2: de slang** *the snake* **V3: zingt** *sings*	**een liedje** a song	**van school** from school
Dutch monolingual	**V2**	149 (250)	206 (201)	196 (202)	156 (183)	323 (316)
	V3	131 (233)	188 (198)	243 (243)	177 (190)	292 (304)
English-Dutch	**V2**	197 (278)	301 (277)	290 (316)	225 (260)	458 (400)
	V3	179 (291)	267 (283)	352 (352)	264 (307)	379 (366)
German-Dutch	**V2**	140 (254)	259 (246)	216 (214)	199 (262)	365 (329)
	V3	159 (287)	244 (270)	297 (305)	223 (262)	325 (338)

Table A4.1.3. Participants' mean residual RTs (and SDs) in ms on the self-paced listening task per segment in the PP-V and V-PP condition for each group in session 2 (bilingual session).

		Segment						
		1 De leeuw *the lion*	2 wordt *is*	3 PP-V:door de beer *by the bear* V-PP: geduwd *pushed*	4 PP-V: geduwd *pushed* V-PP: door de beer *by the bear*	5 en *and*	6 de muis *the mouse*	7 likt (…) *licks*
English-Dutch	PP-V	204 (229)	279 (227)	180 (306)	225 (299)	354 (217)	247 (241)	260 (355)
	V-PP	184 (197)	284 (230)	158 (251)	245 (317)	344 (206)	242 (240)	255 (334)
German-Dutch	PP-V	188 (184)	224 (157)	140 (199)	192 (257)	311 (165)	221 (183)	238 (301)
	V-PP	182 (197)	238 (171)	121 (192)	169 (200)	288 (152)	219 (182)	222 (303)

Table A4.1.4. Participants' mean residual RTs (and SDs) in ms on the self-paced listening task per segment in the V2 and V3 condition for each group in session 2 (bilingual session).

		Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
		op de bank *on the couch*	**V2: zingt** *sings* **V3: de slang** *the snake*	**V2: de slang** *the snake* **V3: zingt** *sings*	**een liedje** a song	**van school** from school
English-Dutch	**V2**	78 (211)	187 (207)	167 (178)	146 (218)	273 (319)
	V3	79 (228)	147 (213)	223 (290)	162 (243)	235 (300)
German-Dutch	**V2**	88 (215)	154 (186)	139 (167)	119 (208)	215 (223)
	V3	64 (183)	125 (181)	165 (208)	128 (181)	207 (269)

A4.2. Model summaries for the analyses of the PP-V and V-PP conditions in session 1 (monolingual session).

Table A4.2.1. Summary of the mixed linear model of English-Dutch, German-Dutch and monolingual Dutch participants' log residual RTs on the long passive sentences (session 1). The reference level is participants' log residual RTs at segment 4 in the PP-V condition.[a]

Group*Segment*Word order + Duration + RT previous trial + Trial number + Forward Digit Span + (1 + Word order + Segment|ID) + (1|Item)							
Random effects		*Var*	*SD*	*Corr*			
ID	Intercept	0.00247	0.04868				
	Word order	0.00013	0.01124	-0.28			
	Segment 3	0.00034	0.01833	0.13	0.23		
	Segment 4	0.00094	0.03064	-0.19	0.31	0.78	
	Segment 5	0.00073	0.02696	-0.48	0.12	-0.05	0.30
Item	Intercept	0.00003	0.00562				
Residual		0.00228	0.04777				
Fixed effects				*B*	*SE*	*t*	*p*
Intercept				3.0773	0.0065	476.4	< .001
Group1 (bilinguals vs. monolinguals)				-0.0216	0.0133	-1.6	.107
Group2 (English-Dutch vs. German-Dutch)				-0.0124	0.0153	-0.8	.421
Segment 1&2				-0.0248	0.0044	-5.7	< .001
Segment 3				-0.0062	0.0032	-1.9	.059
Segment 5				-0.0554	0.0051	-10.8	< .001
Word order				0.0051	0.0025	2.0	.045
Group1 (bilinguals vs. monolinguals)*Word order				0.0161	0.0053	3.0	.003
Group2 (English-Dutch vs. German-Dutch)*Word order				-0.0037	0.0062	-0.6	.548
Group1 (bilinguals vs. monolinguals)*Segment 1&2				0.0121	0.0087	1.4	.171

Table A4.2.1 (continued).

Fixed effects	*B*	*SE*	*t*	*p*
Group1 (bilinguals vs. monolinguals)*Segment 3	0.0170	0.0066	2.6	0.012
Group1 (bilinguals vs. monolinguals)*Segment 5	0.0099	0.0095	1.0	.299
Group2 (English-Dutch vs. German-Dutch)*Segment 1&2	0.0022	0.0101	0.2	.825
Group2 (English-Dutch vs. German-Dutch)*Segment 3	0.0063	0.0078	0.8	.415
Group2 (English-Dutch vs. German-Dutch) * Segment 5	0.0113	0.0110	1.0	.307
Group1 (bilinguals vs. monolinguals)*Segment 1&2*Word order	-0.0203	0.0063	-3.2	.001
Group1 (bilinguals vs. monolinguals)*Segment 3*Word order	-0.0205	0.0063	-3.3	.001
Group1 (bilinguals vs. monolinguals)*Segment 5*Word order	-0.0154	0.0063	-2.4	.014
Group2 (English-Dutch vs. German-Dutch)*Segment 1&2*Word order	0.0065	0.0074	0.9	.379
Group2 (English-Dutch vs. German-Dutch)*Segment 3*Word order	-0.0016	0.0074	-0.2	.833
Group2 (English-Dutch vs. German-Dutch)*Segment 5*Word order	-0.0028	0.0074	-0.4	.702
Segment 1&2*Word order	-0.0008	0.0030	-0.3	.798
Segment 3*Word order	-0.0123	0.0031	-4.0	< .001
Segment 5*Word order	-0.0024	0.0030	-0.8	.420
Duration	-0.00027	0.0000	-36.0	< .001
RT previous trial	0.00005	0.0000	23.4	< .001
Trial number	-0.0017	0.0001	-27.0	< .001
Forward digit span	-0.0017	0.0008	-2.2	.031

[a]Note that differences in the estimates between segments do not directly correspond to differences in participants' residual listening times between segments. This is because the model controls for the effect of *Duration* on children's residual listening times. That is, the longer the *Duration* of the audio fragment, the shorter the corresponding residual listening time.

Table A4.2.2. Summary of the mixed linear model of English-Dutch, German-Dutch and monolingual Dutch participants' log residual RTs on the long passive sentences (session 1). The reference level is participants' log residual RTs at segment 4 in the V-PP condition.[a]

Group*Segment*Word order + Duration + RT previous trial + Trial number + Forward Digit Span + (1 + Word order + Segment|ID) + (1|Item)

Random effects		*Var*	*SD*	*Corr*			
ID	Intercept	0.00247	0.04868				
	Word order	0.00013	0.01124	-0.28			
	Segment 3	0.00034	0.01833	0.13	0.23		
	Segment 4	0.00094	0.03064	-0.19	0.31	0.78	
	Segment 5	0.00073	0.02696	-0.48	0.12	-0.05	0.30
Item	Intercept	0.00003	0.00562				
Residual		0.00228	0.04777				

Fixed effects	*B*	*SE*	*t*	*p*
Intercept	3.0219	0.0057	533.9	<.001
Group1 (bilinguals vs. monolinguals)	-0.0117	0.0110	-1.1	.289
Group2 (English-Dutch vs. German-Dutch)	-0.0011	0.0127	-0.1	.932
Segment 1&2	0.0305	0.0041	7.5	<.001
Segment 3	0.0492	0.0054	9.1	<.001
Segment 5	0.0554	0.0051	10.8	<.001
Word order	0.0027	0.0025	1.1	.277
Group1 (bilinguals vs. monolinguals)*Word order	0.0007	0.0052	0.1	.895
Group2 (English-Dutch vs. German-Dutch)*Word order	-0.0065	0.0060	-1.1	.280
Group1 (bilinguals vs. monolinguals)*Segment 1&2	0.0021	0.0080	0.3	.788

Table A4.2.2 (continued).

Fixed effects	*B*	*SE*	*t*	*p*
Group1 (bilinguals vs. monolinguals)*Segment 3	0.0071	0.0093	0.8	.450
Group1 (bilinguals vs. monolinguals) * Segment 5	-0.0099	0.0095	-1.0	.299
Group2 (English-Dutch vs. German-Dutch)*Segment 1&2	-0.0091	0.0092	-1.0	.327
Group2 (English-Dutch vs. German-Dutch)*Segment 3	-0.0050	0.0108	-0.5	.646
Group2 (English-Dutch vs. German-Dutch)*Segment 5	-0.0113	0.0110	-1.0	.307
Group1 (bilinguals vs. monolinguals)*Segment 1&2*Word order	-0.0049	0.0063	-0.8	.435
Group1 (bilinguals vs. monolinguals)*Segment 3*Word order	-0.0051	0.0062	-0.8	.411
Group1 (bilinguals vs. monolinguals)*Segment 5*Word order	-0.0154	0.0063	2.4	.014
Group2 (English-Dutch vs. German-Dutch)*Segment 1&2*Word order	0.0093	0.0073	1.3	.200
Group2 (English-Dutch vs. German-Dutch)*Segment 3*Word order	0.0013	0.0073	0.2	.863
Group2 (English-Dutch vs. German-Dutch)*Segment 5*Word order	0.0028	0.0074	0.4	.702
Segment 1&2*Word order	0.0017	0.0030	0.6	.571
Segment 3*Word order	-0.0098	0.0030	-3.3	<.001
Segment 5*Word order	0.0025	0.0030	0.8	.420
Duration	-0.00027	0.0000	-36.0	<.001
RT previous trial	0.00005	0.0000	23.4	<.001
Trial number	-0.0017	0.0001	-27.0	<.001
Forward digit span	-0.0017	0.0008	-2.2	.031

[a]Note that differences in the estimates between segments do not directly correspond to differences in participants' residual listening times between segments. This is because the model controls for the effect of *Duration* on children's residual listening times. That is, the longer the *Duration* of the audio fragment, the shorter the corresponding residual listening time.

A4.3. Log likelihood tests for models with main effects of the language dominance measures (*Current listening, Current speaking, Cumulative input* and *Relative proficiency*) and with their interactions with *Word order* and *Segment* by test session by sentence type.

Table A4.3.1. Model improvements after adding main effects of the four dominance measures and their interactions with *Word order* and *Word order* and *Segment* by group for the long passive sentences in session 1 (monolingual session).

	English-Dutch			German-Dutch		
	X^2	*Δdf*	*p*	X^2	*Δdf*	*p*
Current listening English/German (%)	0.4	1	.512	0.4	1	.504
**Word order*	1.0	1	.322	0.8	1	.363
**Word order*Segment*	1.8	2	.404	0.3	2	.882
Current speaking English/German (%)	1.1	1	.287	0.0	1	.878
**Word order*	0.2	1	.687	2.2	1	.137
**Word order*Segment*	1.7	2	.431	1.3	2	.532
Cumulative input English/German (%)	0.0	1	.926	0.7	1	.408
**Word order*	0.5	1	.499	0.2	1	.646
**Word order*Segment*	1.6	2	.440	2.3	2	.323
Relative proficiency	3.0	1	.085	0.1	1	.808
**Word order*	0.9	1	.357	0.6	1	.421
**Word order*Segment*	1.8	2	.410	3.0	2	.218

Table A4.3.2. Model improvements after adding main effects of the four dominance measures and their interactions with *Word order* and *Word order* and *Segment* by group for the V2 and V3 sentences in session 1 (monolingual session).

	English-Dutch			German-Dutch		
	X^2	*Δdf*	*p*	X^2	*Δdf*	*p*
Current listening English/German (%)	1.5	1	.225	0.4	1	.524
**Word order*	0.0	1	.828	1.9	1	.171
**Word order*Segment*	4.0	2	.134	1.4	2	.505
Current speaking English/German (%)	0.4	1	.552	0.4	1	.546
**Word order*	5.8	1	.016	4.1	1	.042
**Word order*Segment*	3.4	2	.180	2.6	2	.272
Cumulative input English/German (%)	0.8	1	.384	1.7	1	.194
**Word order*	1.4	1	.236	0.3	1	.573
**Word order*Segment*	3.0	2	.223	3.6	2	.164
Relative proficiency	3.0	1	.084	0.8	1	.361
**Word order*	0.0	1	.936	1.4	1	.240
**Word order*Segment*	0.4	2	.817	3.0	2	.227

Table A4.3.3. Model improvements after adding main effects of the four dominance measures and their interactions with *Word order* and *Word order* and *Segment* by group for the long passive sentences in session 2 (bilingual session).

	English-Dutch			German-Dutch		
	X^2	*Δdf*	*p*	X^2	*Δdf*	*p*
Current listening English/German (%)	0.3	1	.555	0.4	1	.526
**Word order*	0.0	1	.848	0.3	1	.604
**Word order*Segment*	1.0	2	.599	2.3	2	.322
Current speaking English/German (%)	0.8	1	.368	0.0	1	.866
**Word order*	1.3	1	.249	0.2	1	.626
**Word order*Segment*	5.2	2	.076	0.2	2	.913
Cumulative input English/German (%)	0.0	1	.905	2.0	1	.155
**Word order*	3.5	1	.060	0.0	1	.786
**Word order*Segment*	5.1	2	.077	15.9	2	<.001
Relative proficiency	2.2	1	.138	0.1	1	.742
**Word order*	0.9	1	.352	1.1	1	.302
**Word order*Segment*	0.4	2	.839	0.6	2	.738

Table A4.3.4. Model improvements after adding main effects of the four dominance measures and their interactions with *Word order* and *Word order* and *Segment* by group for the V2 and V3 sentences in session 2 (bilingual session).

	English-Dutch			German-Dutch		
	X^2	*Δdf*	*p*	X^2	*Δdf*	*p*
Current listening English/German (%)	0.0	1	.890	0.3	1	.562
**Word order*	2.5	1	.116	0.0	1	.929
**Word order*Segment*	0.5	2	.793	2.6	2	.270
Current speaking English/German (%)	0.1	1	.811	0.0	1	.933
**Word order*	0.5	1	.471	0.2	1	.681
**Word order*Segment*	0.1	2	.958	4.3	2	.115
Cumulative input English/German (%)	2.4	1	.121	0.8	1	.357
**Word order*	1.1	1	.285	0.9	1	.332
**Word order*Segment*	2.5	2	.292	3.4	2	.181
Relative proficiency	3.0	1	.084	0.1	1	.739
**Word order*	0.0	1	.994	0.0	1	.907
**Word order*Segment*	0.2	2	.925	0.4	2	.831

Appendix Chapter 5

A5.1. List of stimuli and questions used in the eye-tracking task in the optional condition.

Item	Sentence
1	Anna en Lieke leren in de bieb. Terwijl Lieke een boek leest, neemt zij een slokje water. “Anna and Lieke are studying in the library. While Lieke is reading a book, she takes a sip of water.” Wie neemt een slokjes water? “Who takes a sip of water?”
2	Anna en Sophie spelen in de zandbak. Terwijl Anna een zandkasteel bouwt, vult zij een emmer met zand. “Anna and Sophie are playing the sandbox. While Anna is building a sandcastle, she is filling a bucket with sand.” Wie vult een emmer met zand? “Who is filling a bucket with sand?”
3	Sophie en Anna hebben pauze. Terwijl Anna naar muziek luistert, eet zij een boterham met kaas. “Sophie and Anna are having a break. While Anna is listening to music, she is eating a cheese sandwich.” Wie eet een boterham met kaas? “Who is eating a cheese sandwich?”
4	Sophie en Lieke maken muziek op het podium. Terwijl Lieke gitaar speelt, zingt zij een vrolijk liedje. “Sophie and Lieke are playing music on stage. While Lieke is playing guitar, she is singing a cheerful song.” Wie speelt een vrolijk liedje? “Who is playing a cheerful song?”

A5.1 (continued).

Item	Sentence
5	Lieke en Anna zijn bij een muziekconcert. Terwijl Anna op de muziek danst, zingt zij het liedje hardop mee. "Lieke and Anna are at a concert. While Anna is dancing on the music, she is singing the song out loud." Wie zing het liedje hardop mee? "Who is singing the song out loud?"
6	Anna en Sophie kijken naar schilderijen in het museum. Terwijl Anna een schilderij bekijkt, maakt zij een aantal foto's. "Anna and Sophie are looking at paintings at the museum. While Anna is studying a painting, she takes some pictures." Wie maakt een aantal foto's? "Who is taking some pictures?"
7	Sophie en Lieke zijn op het schoolplein. Terwijl Sophie op een bankje gaat zitten, vindt zij een aardbeienlolly. "Sophie and Lieke are at the schoolyard. While Sophie sits down on a bench, she finds a strawberry lollipop." Wie vindt een aardbeienlolly? "Who finds a strawberry lollipop?"
8	Sophie en Anna versieren het klaslokaal voor het schoolfeest. Terwijl Anna slingers ophangt, blaast zij een grote ballon op. "Sophie and Anna are decorating the classroom for the school party. While Anna is hanging garlands, she is blowing up a large balloon." Wie blaast een grote ballon op? "Who is blowing up a large balloon?"

A5.1 (continued).

Item	Sentence
9	Sophie en Lieke kijken een film. Terwijl Sophie chips eet, drinkt zij een beker cola. “Sophie and Lieke are watching a movie. While Sophie is eating crisps, she is drinking a cup of coke.” Wie drinkt een beker cola? “Who is drinking a cup of coke?”
10	Anna en Sophie zitten in de klas. Terwijl Sophie een tekening maakt, kijkt zij naar een som op het bord. “Anna and Sophie are in the classroom. While Sophie is making a drawing, she is looking at an equation at the blackboard.” Wie kijkt naar een som op het bord? “Who is looking at an equation at the blackboard?”
11	Sophie en Anna verven een schuurtje. Terwijl Sophie een appel eet, verft zij de deur donkergroen. “Sophie and Anna are painting a shed. While Sophie is eating an apple, she is painting the door dark green.” Wie verft de deur donkergroen? “Who is painting the door dark green?”
12	Lieke en Sophie lopen door de supermarkt. Terwijl Sophie het winkelwagentje duwt, pakt zij een grote zak snoep. “Lieke and Sophie are walking through the supermarket. While Sophie is pushing the shopping cart, she takes a large bag of candy.” Wie pakt een grote zak snoep? “Who takes a large bag of candy?”

A5.1 (continued).

Item	Sentence
13	Lieke en Sophie varen met een bootje. Terwijl Sophie het bootje bestuurt, zwaait zij naar een andere boot. “Lieke and Sophie are sailing with a boat. While Sophie is steering the boat, she waves at another boat.” Wie zwaait naar een andere boot? “Who waves at another boat?”
14	Lieke en Anna zijn een dagje in de dierentuin. Terwijl Lieke op de plattegrond kijkt, voert zij een nootje aan een aap. “Lieke and Anna are a day at the zoo. While Lieke is looking at the map, she is feeding a monkey a nut.” Wie voert een nootje aan een aap? “Who is feeding a monkey a nut?”
15	Lieke en Anna zijn in het pretpark. Terwijl Lieke in de achtbaan zit, drinkt zij een blikje sinas. “Lieke and Anna are at the theme park. While Lieke is riding the rollercoaster, she is drinking a can of Fanta.” Wie drinkt een blikje sinas? “Who is drinking a can of Fanta?”
16	Lieke en Sophie halen ijsjes bij de ijssalon. Terwijl Lieke bij de kassa betaalt, likt zij aan een bolletje ijs. “Lieke and Sophie are getting ice cream from the parlour. While Lieke is paying at the till, she is licking a scoop of ice cream.” Wie likt aan een bolletje ijs? “Who is licking a scoop of ice cream?”

A5.1 (continued).

Item	Sentence
17	Anna en Lieke rijden mee met de brandweerwagen. Terwijl Anna een helm opzet, zet zij de zwaailichten aan. “Anna and Lieke are riding in the fire truck. While Anna is putting on a helmet, she turns on the flashing lights.” Wie zet de zwaailichten aan? “Who turns on the flashing lights?”
18	Anna en Lieke zetten op de camping een tent op. Terwijl Lieke een luchtbed oppompt, rolt zij de slaapzakken uit. “Anna and Lieke are setting up tent at the camp site. While Lieke is inflating an air mattress, she rolls out the sleeping bags.” Wie rolt de slaapzakken uit? “Who roles out the sleeping bags?”
19	Thomas en Peter eten in een restaurant. Terwijl Peter een hap pizza neemt, leest hij de menukaart door. “Thomas and Peter are eating at a restaurant. While Peter takes a bite of pizza, he is reading the menu.” Wie leest de menukaart door? “Who is reading the menu?”
20	Thomas en Joris maken een kampvuur. Terwijl Thomas hout zoekt, steekt hij een lucifer aan. “Thomas and Joris are making a campfire. While Thomas is looking for wood, he lights a match.” Wie steekt een lucifer aan? “Who lights a match?”

A5.1 (continued).

Item	Sentence
21	Peter en Thomas maken het huiswerk voor rekenen. Terwijl Peter een som voorleest, pakt hij een rekenmachine. "Peter and Thomas are doing the homework for mathematics. While Peter is reading an equation, he takes a calculator." Wie pakt een rekenmachine? "Who takes a calculator?"
22	Joris en Peter spelen een bordspel. Terwijl Joris de dobbelstenen gooit, neemt hij een kaart van de stapel. "Joris and Peter are playing a boardgame. While Joris is throwing the dice, he takes a card from the deck." Wie neemt een kaart van de stapel? "Who takes a card from the deck?"
23	Peter en Thomas houden een kussengevecht in de slaapkamer. Terwijl Peter onder het bed wegduikt, gooit hij met een dik kussen. "Peter and Thomas are having a pillow fight in the bedroom. While Peter dives away underneath the bed, he throws a big pillow." Wie gooit met een dik kussen? "Who throws a big pillow?"
24	Thomas en Peter werken in de tuin. Terwijl Peter de planten water geeft, trekt hij het onkruid uit de grond. "Thomas and Peter are working in the garden. While Peter is watering the plants, he is pulling the weeds out of the ground." Wie trekt het onkruid uit de grond? "Who is pulling the weeds out of the ground?"

A5.1 (continued).

Item	Sentence
25	Joris en Peter liggen op het strand. Terwijl Joris zonnebrandcrème pakt, zet hij een vrolijk petje op. “Joris and Peter are lying on the beach. While Joris takes sunscreen, he puts up a cheerful cap.” Wie zet een vrolijk petje op? “Who puts on a cheerful cap?”
26	Joris en Thomas gaan de eenden voeren in het park. Terwijl Thomas brood pakt, lokt hij een groepje eenden. “Joris and Thomas are going to feed the ducks in the park. While Thomas takes bread, he is luring a group of ducks.” Wie lokt een groepje eenden? “Who is luring a group of ducks?”
27	Thomas en Joris gaan de hond uitlaten. Terwijl Thomas de hondenriem vasthoudt, doet hij de voordeur open. “Thomas and Joris are going to walk the dog. While Thomas is holding the dog leash, he opens the front door.” Wie doet de voordeur open? “Who opens the front door?”
28	Joris en Thomas ruimen het klaslokaal op. Terwijl Thomas alle boeken opstapelt, veegt hij de vloer met een bezem. “Joris and Thomas are cleaning up the classroom. While Thomas is piling all the books, he is sweeping the floor with a broom.” Wie veegt de vloer met een bezem? “Who is sweeping the floor with a broom?”

A5.1 (continued).

Item	Sentence
29	Peter en Joris hebben gym in de gymzaal. Terwijl Joris bij de trampoline staat, maakt hij een dubbele salto. “Peter and Joris are having gym in the gym. While Joris is standing at the trampoline, he does a double somersault.” Wie maakt een dubbele salto? “Who does a double somersault?”
30	Joris en Thomas helpen met verhuizen. Terwijl Thomas een doos tilt, eet hij een broodje met ham. “Joris and Thomas are helping with moving. While Thomas is carrying a box, he is eating a ham sandwich.” Wie eet een broodje met ham? “Who is eating a ham sandwich?”
31	Peter en Thomas zijn uit logeren. Terwijl Peter een pyjama aantrekt, leest hij een superspannend boek. “Peter and Thomas are at a sleepover. While Peter puts on pyjamas, he is reading a super exciting book.” Wie leest een superspannend boek? “Who is reading a super exciting book?”
32	Peter en Joris staan op de trein te wachten. Terwijl Joris een flesje water koopt, houdt hij de tijd in de gaten. “Peter and Joris are waiting for the train. While Joris is buying a bottle of water, he is keeping an eye on the time.” Wie houdt de tijd in de gaten? “Who is keeping an eye on the time?”

A5.1 (continued).

Item	Sentence
33	Peter en Joris spelen basketbal. Terwijl Joris het doel verdedigt, gooit hij de basketbal weg. Peter and Joris are playing basketball. While Joris is defending the goal, he throws away the basketball.” Wie gooit de basketbal weg? “Who throws away the basketball?”
34	Thomas en Peter zitten te vissen. Terwijl Peter een schepnet pakt, vist hij een vis uit het water. “Thomas and Peter are fishing. While Peter gets a landing net, he fishes a fish from the water.” Wie vist een vis uit het water? “Who fishes a fish from the water?”
35	Thomas en Joris zijn in de speeltuin. Terwijl Thomas snoepjes eet, klimt hij de hoge glijbaan op. “Thomas and Joris are at the play yard. While Thomas is eating candy, he is climbing onto the high slide.” Wie klimt de hoge glijbaan op? “Who is climbing onto the high slide?”
36	Joris en Peter zijn op de kinderboerderij. Terwijl Joris naar de varkens kijkt, speelt hij met een klein geitje. “Joris and Peter are at the petting farm. While Joris is watching the pigs, he is playing with a little goat.” Wie speelt met een klein geitje? “Who is playing with a little goat?”

A5.2. Description of the Turkish pronoun task: goal, method, results and conclusions.

Goal

The goal of the Turkish pronoun task was to establish whether Turkish-Dutch children have acquired the discourse properties of null and overt pronouns in Turkish. We expected children to link null pronouns more often to the topic of the discourse in Turkish than overt pronouns.

Method

Materials

To assess Turkish-Dutch children's pronoun comprehension in Turkish, we created a Turkish version of the (comprehension question part of the) Dutch task. The structure of the Turkish stories was similar to the Dutch stories. However, because Turkish does not have grammatical gender, we created three different versions of the stories: (i) one with a null pronoun; (ii) one with the pronoun "o"; and (iii) one with the pronoun "o" accompanied by the emphasis marker "da" (see Table A5.2.1). The expected response in the null-pronoun condition was the local interpretation. The expected response in the overt pronoun conditions was the disjoint interpretation.

Apart from the pronouns, Turkish items were a direct translation of the Dutch items to keep the tasks as similar as possible, except for the following changes: (i) the Dutch characters were replaced with typical Turkish names (Zeynep, Leyla, Ayşe, Yusuf, Osman and Caner); (ii) when direct translations of the Dutch sentences were resulted in semantically incorrect sentences, the content was modified accordingly; (iii) Turkish items were not aligned in length of duration, because no eye-tracking was involved.

The set-up of the Turkish tasks was similar to the set-up of the Dutch task. Hence, children listened to 36 pseudo-randomized experimental items (one version of each story) interspersed with 36 filler items. Each item was accompanied by three pictures on the screen: two characters left and right and one object below. After each item, a question was asked. In case of the experimental items, the question served to probe children's interpretation of the pronoun. Sentences were recorded by female native speakers of Turkish using neutral intonation.

Table A5.2.1. Example of a story used in the Turkish comprehension task in the three conditions: optional, local and disjoint.

	Null pronoun	'o' pronoun	'o da' pronoun
Context sentence	<u>Zeynep</u> ve <u>Yusuf</u> kütüphanede çalışıyorlar. *Zeynep and Yusuf are studying in the library.*	<u>Yusuf</u> ve <u>Osman</u> kütüphanede çalışıyorlar. *Yusuf and Osman are studying in the library.*	<u>Zeynep</u> ve <u>Leyla</u> kütüphanede çalışıyorlar. *Zeynep and Leyla are studying in the library.*
Subordinate clause	<u>Yusuf</u> kitap okurken, *While Yusuf is reading a book,*	<u>Yusuf</u> kitap okurken, *While Yusuf is reading a book,*	<u>Zeynep</u> kitap okurken, *While Zeynep is reading a book,*
Main clause	<u>Ø</u> sudan bir yudum alıyor. *<u>Ø</u> takes a sip of water.*	<u>o</u> sudan bir yudum alıyor. *<u>she/he</u> takes a sip of water.*	<u>o da</u> sudan bir yudum alıyor. *<u>she/he</u> takes a sip of water.*
Question	Kim sudan bir yudum alıyor? *Who takes a sip of water?*	Kim sudan bir yudum alıyor? *Who takes a sip of water?*	Kim sudan bir yudum alıyor? *Who takes a sip of water?*

Procedure

The procedure for the Turkish and German pronoun comprehension tasks was similar to the procedure for the Dutch task. The only exception was that no eye-tracking data were recorded.

Scoring and data preparation

Button responses were coded as 1 (local response) or 0 (disjoint response) for each item.

Results

Descriptives

Figure A5.2.1 shows bilingual children's pronoun interpretations in the Turkish task. Overall children had a slight preference for the local interpretation in the null pronoun condition (m = 58.3, SD = 40.0). Children's percentage of local choices was smaller in the overt pronoun conditions (*o*: m = 46.9, SD = 41.7; *o da*: m = 51.0; SD = 44.4).

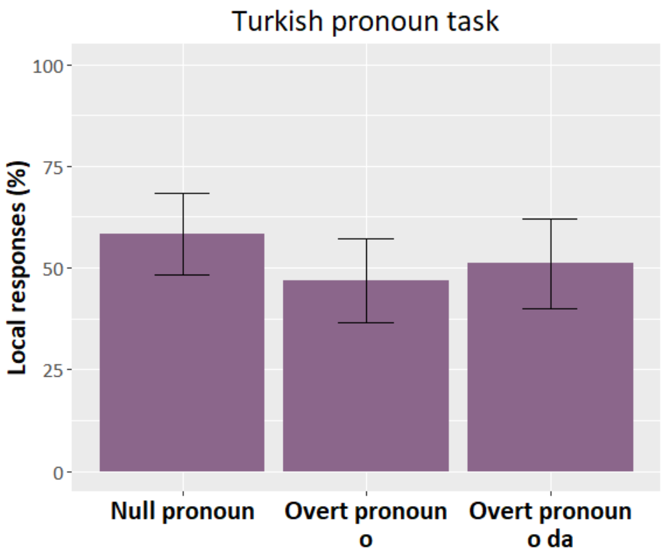


Figure A5.2.1. Average percentage of choices for the local referent on the Turkish pronoun task by *condition*. Error bars represent standard errors.

Data analysis

We ran generalized lmms to test for the effect of condition with random intercepts by child and item. Steps were similar to the analyses of the offline

Dutch data. First, we created a base model with random intercepts by child and item, a fixed effect of *trial number* and significant fixed effects of our background variables (*age, digit span forward, digit span backward* and *Dutch CLT score*). Second, we added the main effect of *condition* which were Helmert contrast coded (contrast 1: disjoint = 2/3, local = -1/3, and optional = -1/3; contrast 2: disjoint = 0, local = 1/2, optional = -1/2). The significance of effects was tested by comparing the fit of models with and without the effect of interest using likelihood ratio tests. Where possible, model stress was reduced by removing absolute model residuals above 2.5.

The main effect of condition was significant (X^2 = 37.0; Δdf = 2; p < .001). The model summary with Helmert contrasts (Table 5.2.2) showed that the Turkish-Dutch children chose the local referent significantly more often as the antecedent of a null pronoun than as the antecedent of an overt pronoun. Children were also more likely to choose the local referent in the *o da* condition compared to the *o* condition. This difference approached significance.

Table 5.2.2. Summary of the generalized linear mixed model of children's local referent choices (percentage) with the main effect of *condition*.

Random effects	*Variance*		*SD*	
ID (intercept)	70.9		8.4	
item (intercept)	1.8		1.4	
Fixed effects	*B*	*SE*	*z*	*p*
intercept	0.96	2.52	0.4	.704
trial number	0.02	0.02	0.7	.455
condition contrast 1 (null vs. overt)	2.80	0.59	4.7	<.001
condition contrast 2 (*o* vs. *o da*)	1.06	0.60	1.76	.079

Conclusion

In sum, Turkish-Dutch children did not show a clear preference for the local interpretation with null pronouns or for a disjoint interpretation with overt pronouns in Turkish: Children's referent choices were around chance level in all conditions. At the same time, however, children showed a greater preference for a local interpretation with a null pronoun than with an overt pronoun. This suggests that the Turkish-Dutch children in our study were sensitive to the different properties of null and overt pronouns in Turkish. In other words, our assumption that Turkish monolingual children associate a Turkish overt pronoun with a non-topic antecedent was borne out, at least in comparison to the Turkish null pronoun. This Turkish association could thus have affected children's pronoun resolution in Dutch as we predicted.

A5.3 Summaries of the linear mixed models and generalized additive mixed models of the eye-tracking data.

Table A5.3.1. Summaries of the separate linear mixed models of children's fixations with the main effect of *time window* (model 1) and the interaction between *time window* and *group* (model 2) and the generalized additive mixed model with the smooth of *time* by *group* (model 3) in the local condition. The pronoun time window level is the reference level in all models. In model 3 the group reference level is the Dutch monolingual group.

Local condition

	Random effects	Variance	*SD*	Fixed effects	*B*	*SE*	*t*	*p*
Model 1 (lmer)	ID (intercept)	0.53	0.726	intercept	-0.96	0.132	-7.2	<.001
	item (intercept)	0.23	0.480	trial number	0.01	0.002	5.5	<.001
	residual	11.08	3.329	time window	2.21	0.041	53.5	<.001
Model 2 (lmer)	ID (intercept)	0.52	0.133	intercept	-1.03	0.113	-7.7	<.001
	item (intercept)	0.23	0.482	trial number	0.01	0.002	5.5	<.001
	residual	11.06	3.326	time window	2.28	0.042	53.8	<.001
				group contrast 1 (Turkish-Dutch vs. others)	-0.07	0.226	-0.3	.747
				group contrast 2 (German-Dutch vs. Dutch)	0.86	0.259	3.3	.002
				time window : group contrast 1	0.07	0.091	0.8	.446
				time window : group contrast 2	-0.69	0.103	-6.8	<.001

	Smooth terms	*edf*	*Ref.df*	*F*	*p*	Fixed effects	*B*	*SE*	*t*	*p*
Model 3 (gamm)	s(time)	11.9	14.8	4.7	<.001	intercept	0.31	0.303	1.0	.305
	s(time) : Dutch monolingual	9.6	12.6	1.7	.048	trial number	0.32	0.004	8.5	<.001
	s(time) : German-Dutch	1.0	1.0	0.7	.418	Dutch CLT score	-0.06	0.021	-2.8	.005
	s(time) : Turkish-Dutch	4.4	5.9	1.6	.137	German-Dutch group	0.593	.368	1.6	.107
	s(ID)	32.0	49.0	2.7	<.001	Turkish-Dutch group	-0.41	0.416	-1.0	.329
	s(item)	28.2	35.0	6.4	<.001					

Table 5.3.2. Summaries of the separate linear mixed models of Turkish-Dutch children's fixations in the local condition with significant main effects of and interactions with the three dominance measures (*relative proficiency*, *current input* and *cumulative input*). The pronoun time window level is the reference level in all models.

	Local condition							
	Random effects	Variance	*SD*	Fixed effects	*B*	*SE*	*t*	*p*
Relative proficiency * time window	ID (intercept)	0.85	0.923	intercept	-0.36	0.328	-1.1	.283
	item (intercept)	0.69	0.828	trial number	0.01	0.004	2.6	.009
	residual	11.42	3.379	time window	1.86	0.100	18.7	<.001
				relative proficiency	0.03	0.010	3.3	.005
				time window:relative proficiency	-0.02	0.004	-5.8	<.001
Current input * time window	ID (intercept)	0.84	0.919	intercept	-2.00	0.870	-2.3	.035
	item (intercept)	0.86	0.925	trial number	0.01	0.004	2.5	.013
	residual	11.47	3.386	time window	2.98	0.293	10.2	<.001
				current input	0.04	0.028	1.3	.227
				time window:current input	-0.03	0.009	-2.7	.007
Cumulative input * time window	ID (intercept)	0.85	0.924	intercept	-0.37	0.776	-0.5	.641
	item (intercept)	0.83	0.914	trial number	0.01	0.004	2.7	.007
	residual	11.40	3.377	time window	0.29	0.296	1.0	.321
				cumulative input	-0.01	0.018	-0.7	.488
				time window:cumulative input	0.05	0.007	6.8	<.001

Table 5.3.3. Summaries of the separate linear mixed models of children's fixations with the main effect of *time window* (model 1) and the interaction between *time window* and *group* (model 2) and the generalized additive mixed model with the smooth of *time* by *group* (model 3) in the disjoint condition. The pronoun time window level is the reference level in all models. In model 3 the group reference level is the Dutch monolingual group.

	Disjoint condition									
	Random effects	Variance	*SD*			Fixed effects	*B*	*SE*	*t*	*p*
Model 1	ID (intercept)	1.16	1.077			intercept	-0.74	0.181	-4.1	<.001
(lmer)	item (intercept)	0.35	0.591			trial number	-0.01	0.002	-4.5	<.001
	residual	12.19	3.492			time window	-0.24	0.043	-5.5	<.001
Model 2	ID (intercept)	1.15	1.074			intercept	-0.73	0.183	-4.0	<.001
(lmer)	item (intercept)	0.35	0.590			trial number	-0.01	0.002	-4.4	<.001
	residual	12.18	3.490			time window	-0.22	0.045	-5.0	<.001
						group contrast 1 (Turkish-Dutch vs. others)	-0.09	0.329	-0.3	.796
						group contrast 2 (German-Dutch vs. Dutch)	-0.14	0.378	-0.4	.715
						time window:group contrast 1	-0.51	0.095	-5.3	<.001
						time window:group contrast 2	-0.28	0.108	-2.6	.010
	Smooth terms	edf	*Ref.df*	*F*	*p*	Fixed effects	*B*	*SE*	*t*	*p*
Model 3	s(time)	1.0	1.0	5.0	.026	intercept	-0.76	0.410	-1.8	.066
(gamm)	s(time):Dutch monolingual	0.0	0.0	0.1	.986	trial number	0.00	0.004	0.1	.929
	s(time):German-Dutch	5.6	7.6	10.7	<.001	digit span backward	-0.12	0.056	-2.1	.033
	s(time):Turkish-Dutch	1.0	1.0	21.6	<.001	German-Dutch group	-0.17	0.497	-0.3	.734
	s(ID)	38.5	49.0	4.8	<.001	Turkish-Dutch group	-0.09	0.522	-0.2	.864
	s(item)	29.6	35.0	9.6	<.001					

Table 5.3.4. Summaries of the separate linear mixed models of Turkish-Dutch children's fixations in the disjoint condition with significant main effects of and interactions with the three dominance measures (*relative proficiency*, *current input* and *cumulative input*). The pronoun time window level is the reference level in all models.

	Disjoint condition							
	Random effects	Variance	*SD*	Fixed effects	*B*	*SE*	*t*	*p*
Relative proficiency	ID (intercept)	0.65	0.805	intercept	-0.52	0.370	-1.4	.175
	item (intercept)	1.05	1.023	trial number	-0.03	0.004	-6.2	<.001
	residual	12.51	3.537	relative proficiency	0.03	0.011	2.7	.016
Current input * time window	ID (intercept)	0.65	0.805	intercept	-1.41	1.073	-1.3	.208
	item (intercept)	1.34	1.159	trial number	-0.03	0.004	-6.2	<.001
	residual	12.37	3.517	time window	-2.00	0.302	-6.6	<.001
				current input	0.03	0.035	0.7	.478
				time window:current input	0.04	0.010	4.5	<.001
Cumulative input * time window	ID (intercept)	0.65	0.806	intercept	-2.33	0.958	-2.4	.027
	item (intercept)	1.61	1.268	trial number	-0.02	0.004	-6.2	<.001
	residual	12.36	3.52	time window	0.76	0.290	2.6	.009
				cumulative input	0.43	0.247	1.7	.103
				time window:cumulative input	-0.37	0.071	-5.2	<.001

Table 5.3.5. Summaries of the separate linear mixed models of children's fixations with the main effect of *time window* (model 1) and the interaction between *time window* and *group* (model 2) and the generalized additive mixed model with the smooth of *time* by *group* (model 3) in the optional condition. The pronoun time window level is the reference level in all models. In model 3 the group reference level is the Dutch monolingual group.

	Optional condition							
	Random effects	**Variance**	***SD***	**Fixed effects**	***B***	***SE***	***t***	***p***
Model 1	ID (intercept)	1.04	1.022	intercept	-0.55	0.169	-3.3	.002
(lmer)	item (intercept)	0.27	0.524	trial number	-0.01	0.002	-2.2	.027
	residual	11.95	3.457	time window	1.00	0.043	23.5	<.001
Model 2	ID (intercept)	1.09	1.043	intercept	-0.57	0.174	-3.3	.002
(lmer)	item (intercept)	0.27	0.523	trial number	-0.01	0.002	-2.2	.026
	residual	11.95	3.46	time window	1.03	0.044	23.3	<.001
				group contrast 1 (Turkish-Dutch vs. others)	-0.25	0.319	-0.8	.429
				group contrast 2 (German-Dutch vs. Dutch)	0.09	0.367	0.3	.800
				time window:group contrast 1	0.26	0.094	2.8	.006
				time window:group contrast 2	-0.16	0.107	-1.5	.127

	Smooth terms	**edf**	***Ref.df***	***F***	***p***	**Fixed effects**	***B***	***SE***	***t***	***p***
Model 3	s(time)	10.1	13.2	8.3	<.001	intercept	0.12	0.425	0.3	.782
(gamm)	s(time):Dutch monolingual	0.0	0.0	0.0	.996	trial number	0.02	0.004	5.1	<.001
	s(time):German-Dutch	1.0	1.0	7.0	.008	German-Dutch group	-0.09	0.510	-0.2	.863
	s(time):Turkish-Dutch	1.0	1.0	7.6	.006	Turkish-Dutch group	-0.13	.538	-0.2	.813
	s(ID)	40.4	50.0	5.8	<.001					
	s(item)	30.7	35.0	16.0	<.001					

Table 5.3.6. Summaries of the separate linear mixed models of Turkish-Dutch children's fixations in the optional condition with significant main effects of and interactions with the three dominance measures (*relative proficiency*, *current input* and *cumulative input*). The pronoun time window level is the reference level in all models.

Optional condition					
	Random effects	*Var*		*SD*	
Relative proficiency	ID (intercept)	0.84		0.915	
	item (intercept)	1.35		1.164	
	residual	12.23		3.50	
	Fixed effects	*B*	*SE*	*t*	*p*
	intercept	0.66	0.420	1.6	.134
	trial number	-0.00	0.004	-0.7	.499
	relative proficiency	0.03	0.013	2.3	.035

A5.4. Figures of the generalized additive mixed models of the eye-tracking data in the local and disjoint condition.

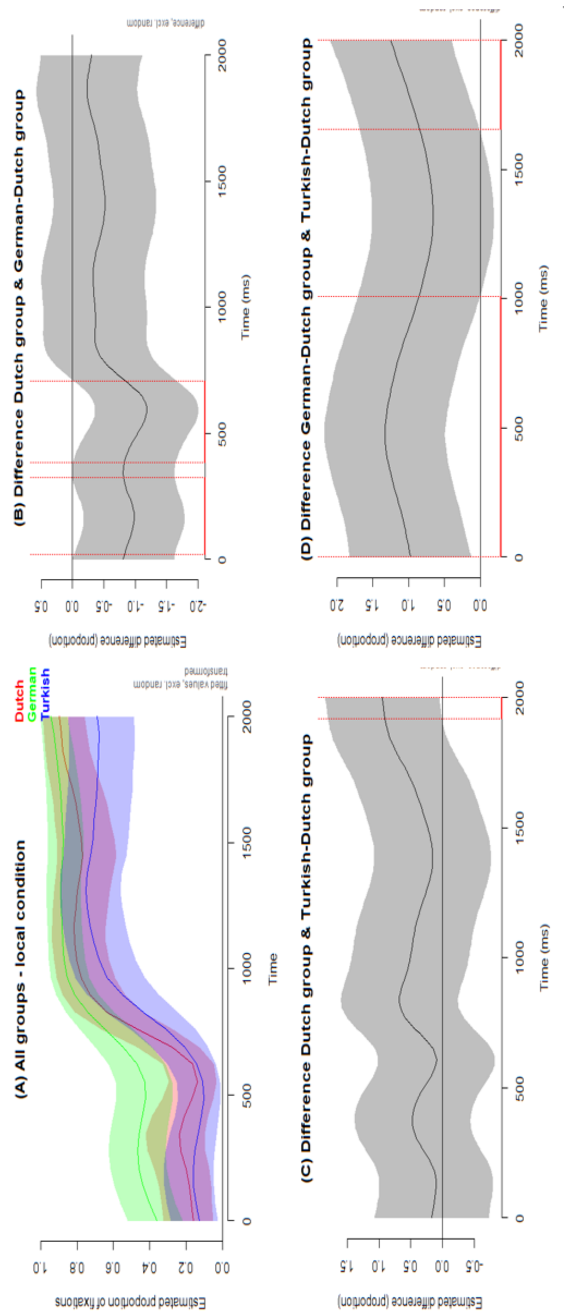


Figure 5.4.1. Panel A: Estimated difference in fixations (back-transformed to proportions) over *time* by *group* in the local condition with 95% confidence intervals. Panel B-D: Estimated difference in difference fixations between groups over *time* with 95% confidence intervals. The areas between the vertical red lines indicate significant differences between groups.

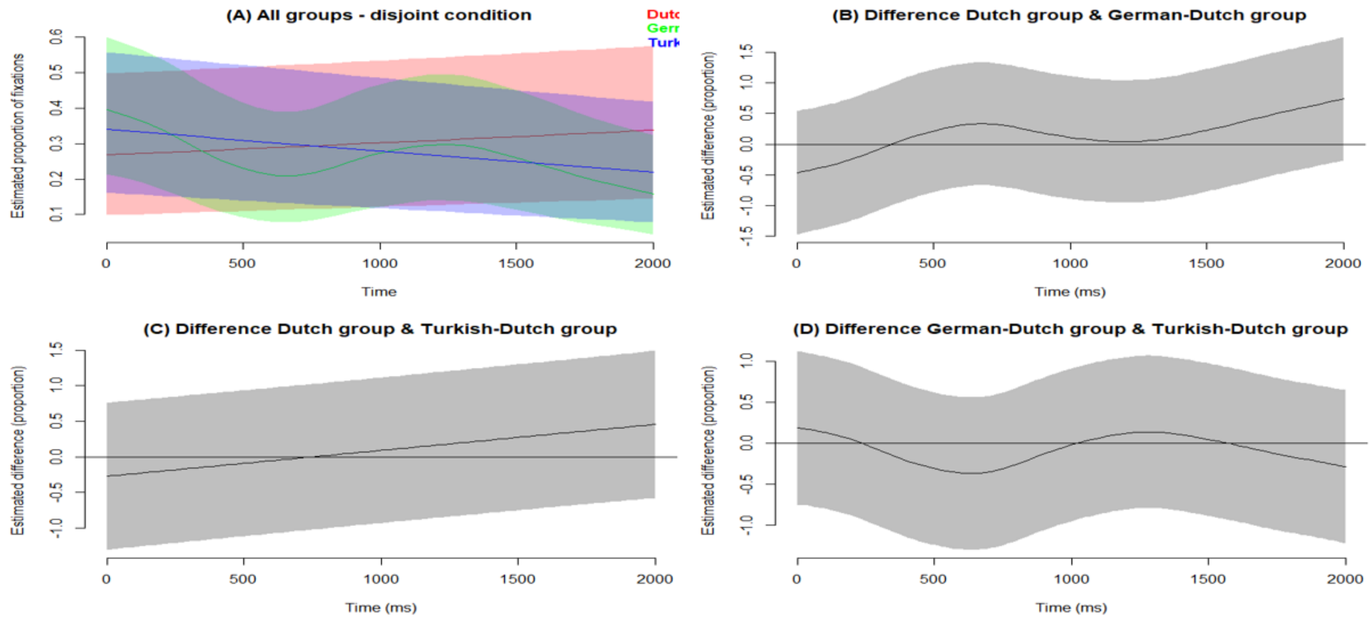


Figure 5.4.2. Panel A: Estimated difference in fixations (back-transformed to proportions) over *time* by *group* in the disjoint condition with 95% confidence intervals. Panel B-D: Estimated difference in difference fixations between groups over *time* with 95% confidence intervals. The areas between the vertical red lines indicate significant differences between groups.

A5.5. Figures of Turkish-Dutch children's average proportion of fixations on the local and disjoint referent in the local, disjoint and optional condition by *current input* and *cumulative input*.

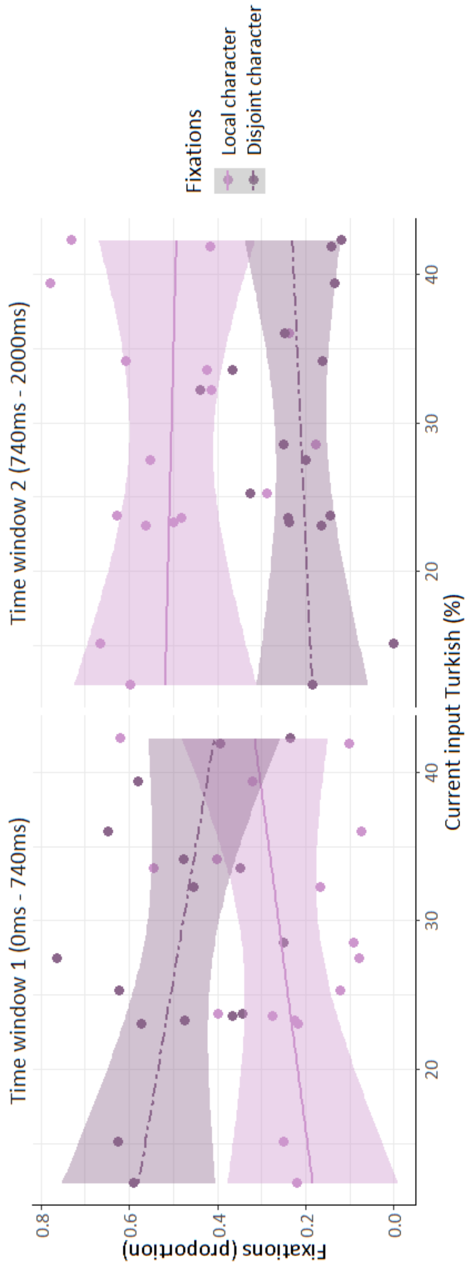


Figure 5.5.1. Turkish-Dutch children's average proportions of fixations on the local and disjoint referent in the local condition by *time window* by *current input* (Turkish) with 95% confidence intervals.

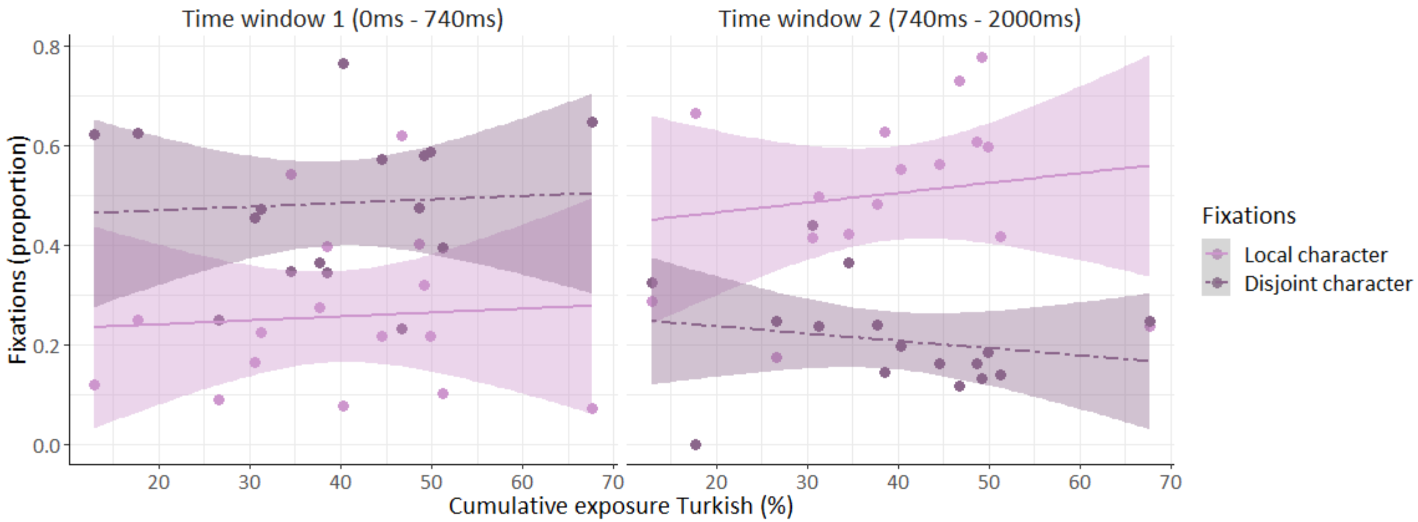


Figure 5.5.2. Turkish-Dutch children's average proportions of fixations on the local and disjoint referent in the local condition by *time window* by *cumulative input* (Turkish) with 95% confidence intervals.

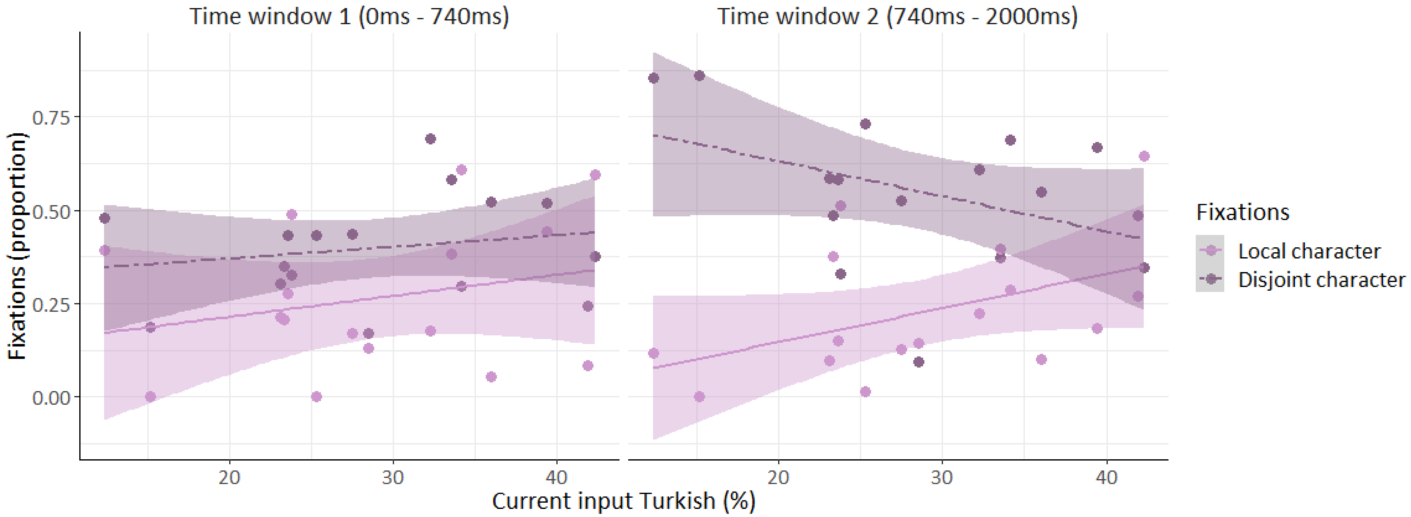


Figure 5.5.3. Turkish-Dutch children's average proportions of fixations on the local and disjoint referent in the disjoint condition by *time window* by *current input* (Turkish) with 95% confidence intervals.

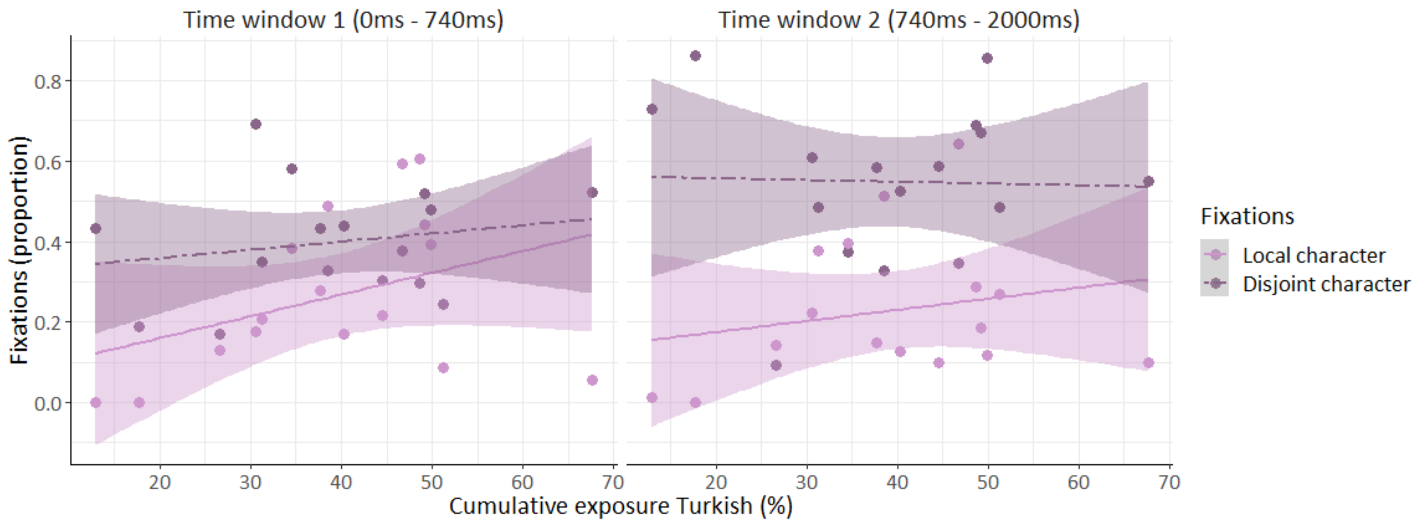


Figure 5.5.4. Turkish-Dutch children's average proportions of fixations on the local and disjoint referent in the disjoint condition by *time window* by *cumulative input* (Turkish) with 95% confidence intervals.

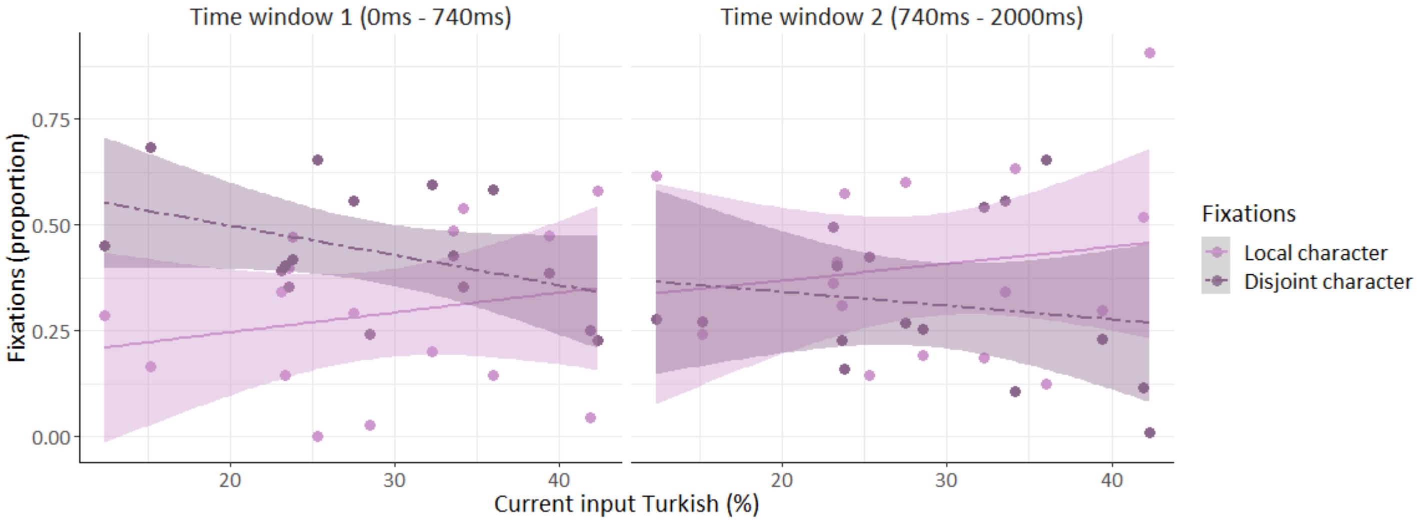


Figure 5.5.5. Turkish-Dutch children's average proportions of fixations on the local and disjoint referent in the optional condition by *time window* by *current input* (Turkish) with 95% confidence intervals.

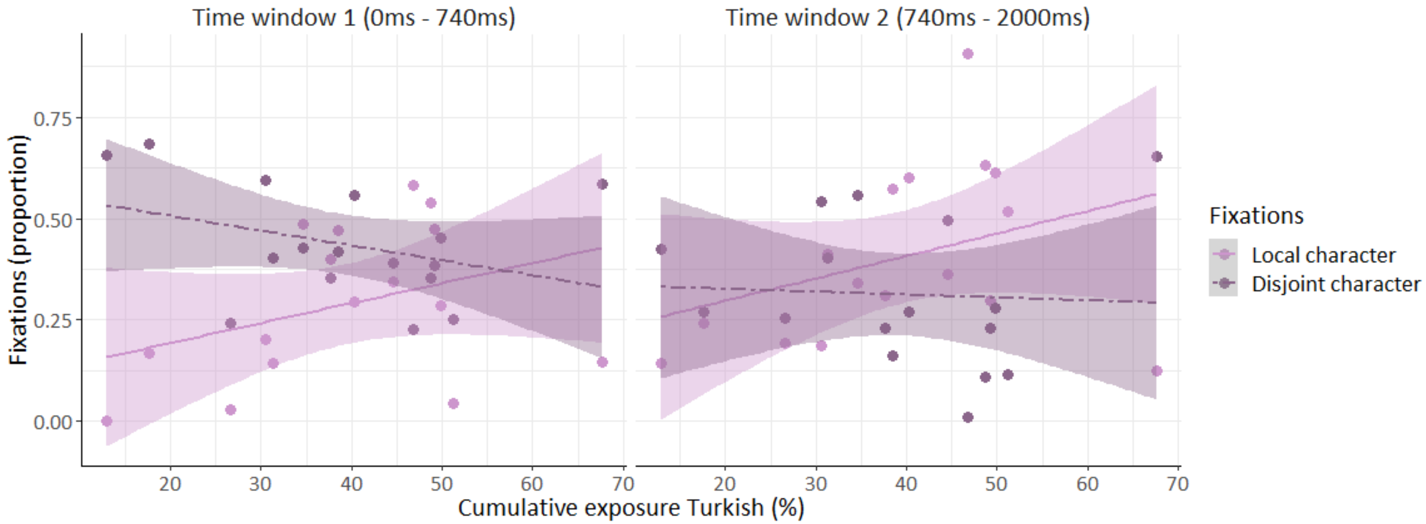


Figure 5.5.6. Turkish-Dutch children's average proportions of fixations on the local and disjoint referent in the optional condition by *time window* by *cumulative input* (Turkish) with 95% confidence intervals.

A5.6. Means and standard deviations of children's offline pronoun choices.

Table 5.6.1. Means and standard deviations of children's offline choices for the local referent (percentages) by *condition* by *group*.

		Dutch group		German-Dutch group		Turkish-Dutch group	
		m	*SD*	*m*	*SD*	*m*	*SD*
Condition	Local	86.9	18.4	86.0	31.9	74.5	31.8
	Disjoint	28.6	40.9	28.0	37.6	15.7	30.7
	Optional	57.7	20.0	75.8	27.4	57.4	31.6

A5.7. Summaries of the generalized linear mixed models of children's offline pronoun resolution data.

Table 5.7.1. Summaries of the separate generalized linear mixed models of children's local referent choices (percentage) with the main effect of *condition* (model 1) and the interaction between *condition* and *group* (model 2). In model 2 the disjoint condition is at the reference level.

	Offline pronoun choices							
	Random effects	Variance	*SD*	Fixed effects	*B*	*SE*	*z*	*p*
Model 1	ID (intercept)	13.7	3.7	intercept	0.88	0.53	1.7	.092
				trial number	-0.01	0.01	-1.3	.198
				condition contrast 1 (disjoint vs. local & optional)	-5.63	0.33	-16.9	<.001
				condition contrast 2 (local vs. optional)	1.71	0.20	8.5	<.001
Model 2	ID (intercept)	5.0	2.2	intercept	-1.89	0.34	-5.5	<.001
				trial number	-0.01	0.01	-0.9	.393
				local condition	4.47	0.23	19.4	<.001
				optional condition	2.98	0.20	15.0	<.001
				group contrast 1 (Turkish-Dutch vs. others)	-1.16	0.74	-1.56	.118
				group contrast 2 (German-Dutch vs. Dutch)	-0.46	0.83	-0.6	.574
				local condition:group contrast 1	-0.01	0.48	-0.0	.977
				local condition:group contrast 2	1.18	0.54	2.2	.031
				optional condition:group contrast 1	0.43	0.43	1.0	.313
				optional condition:group contrast 2	2.00	0.45	4.4	<.001

Table 5.7.2. Summary of the generalized linear mixed model with Turkish-Dutch children's offline local referent choices (percentage) with the interaction between *relative proficiency* and *condition* including all children. The disjoint condition is at the reference level.

	Random effects	*Var*		*SD*	
Relative proficiency *condition (all children)	ID (intercept)	3.8		2.0	
	Fixed effects	*B*	*SE*	*t*	*p*
	intercept	-1.34	0.71	-1.9	.061
	trial number	-0.04	0.01	-3.0	.002
	local condition	3.70	0.47	7.9	<.001
	optional condition	2.73	0.43	6.4	<.001
	relative proficiency	0.09	0.03	3.0	.003
	relative proficiency:local condition	-0.05	0.02	-2.4	.017
	relative proficiency:optional condition	-0.05	0.02	-2.4	.017

A5.8. Figures of Turkish-Dutch children's average proportions of local referent choices (choices) by their *current input* and *cumulative input* in Turkish.

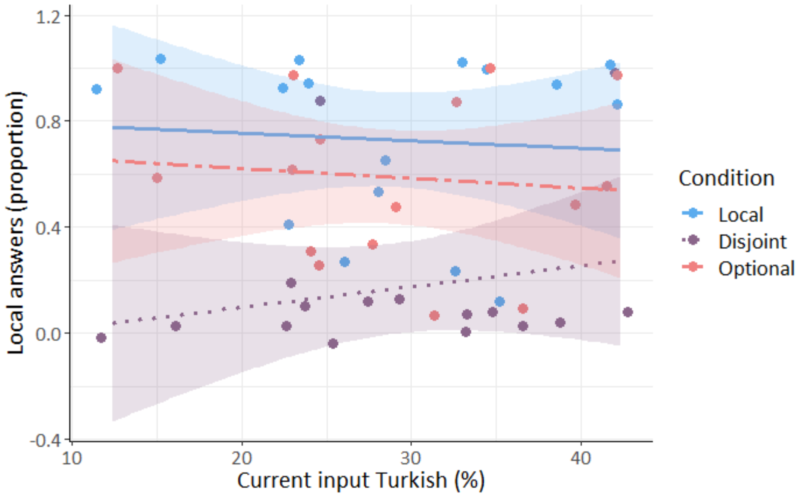


Figure 5.8.1. Turkish-Dutch children's average proportions of local referent choices by *condition* and *current input* (Turkish) with 95% confidence intervals.

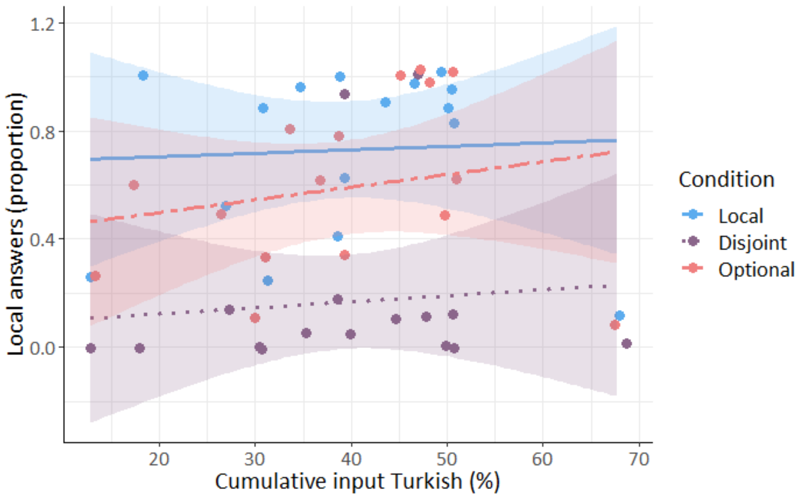


Figure 5.8.2. Turkish-Dutch children's average proportions of local referent choices by *condition* and *cumulative input* (Turkish) with 95% confidence intervals.

Nederlandse samenvatting

Hoe de talen van tweetalige kinderen elkaar beïnvloeden

Tweetalige kinderen zijn goed in staat hun talen uit elkaar te houden (Paradis & Genesee, 1996). Als je ze in hun ene taal hoort praten, merk je vaak helemaal niet dat ze eigenlijk nog een andere taal spreken. Toch gebeurt het af en toe dat de ene taal de andere taal beïnvloedt. Soms gebruiken kinderen een woord van de ene taal in de andere taal (*code-mixing*). Maar beïnvloeding tussen talen kan ook subtieler zijn. Dat is bijvoorbeeld het geval wanneer kinderen een woordvolgorde van de ene taal in de andere taal gebruiken, zonder de woorden van de talen te mixen. Een Frans-Nederlands tweetalig kind zou bijvoorbeeld kunnen zeggen: *Waarom je huilt?* (Strik & Pérez-Leroux, 2011, p. 194) of *Dat is een kat boos* (gebaseerd op Nicoladis, 2006). Prima zinsstructuren in het Frans, maar beide ongrammaticaal in het Nederlands. Zulke situaties waarin kinderen alleen een eigenschap van de zinsstructuur (syntax) of woordstructuur (morfologie) van de ene taal in de andere taal gebruiken, wordt in de literatuur *cross-linguistic influence* genoemd, letterlijk vertaald *tussen-talige invloed*. Dit is het thema van dit proefschrift.

Aan de ene kant weten we behoorlijk veel over tussen-talige invloed in de taal*productie* van kinderen (voor een overzicht, zie Serratrice, 2013). We zien dat tussen-talige invloed voorkomt in verschillende taalcombinaties, tussen nauw verwante talen, zoals het Frans en Engels (Nicoladis, 2006), maar ook tussen niet-verwante talen, zoals het Kantonees en Engels (Kidd, Chan & Chiu, 2015). Soms leidt tussen-talige invloed tot *kwalitatieve* verschillen tussen tweetalige en eentalige kinderen. Dit betekent dat een tweetalig kind iets op een andere manier zegt dan een eentalig kind dat zou doen. Een voorbeeld daarvan is de woordvolgorde van de zin *Waarom je huilt?* in het Nederlands (Strik & Pérez-Leroux, 2011). Meestal leidt tussen-talige invloed echter tot *kwantitatieve* verschillen tussen tweetalige en eentalige kinderen. Tweetalige kinderen zeggen dan iets wat eentalige kinderen ook zouden kunnen zeggen, maar doen dit, onder invloed van hun andere taal, vaker. Een Engels-Nederlands tweetalig kind gebruikt waarschijnlijk vaker een zin zoals *de fiets van Els* ten opzichte van *Els haar fiets*, omdat de eerste woordvolgorde veel in het Engels voorkomt (op basis van Nicoladis, 2012). Tegelijkertijd lijkt het erop dat tussen-talige invloed vaker of zelfs alleen voorkomt onder bepaalde omstandigheden. In dit proefschrift onderzoeken we drie van deze omstandigheden, namelijk de overlap tussen de talen, de taaldominantie en de leeftijd van het kind. In de volgende sectie leg ik uit wat deze factoren inhouden.

Aan de andere kant weten we nog heel weinig over tussen-talige invloed in het taal*begrip* van kinderen. Weinig studies hebben hier onderzoek naar gedaan (Kidd et al., 2015; Serratrice, 2007). Een uitzondering is bijvoorbeeld een studie van Ludovica Serratrice (2007). Zij onderzocht of kennis van het Engels het interpreteren van persoonlijke voornaamwoorden in het Italiaans beïnvloedt, wanneer Italiaans-Engels tweetalige kinderen zinnen horen zoals in (1).

(1) De portier begroet de postbode, terwijl **hij** de deur opendoet.

In het Engels verwijst *hij* bij voorkeur naar de portier. Maar in het Italiaans verwijst *hij* bij voorkeur naar de postbode. In het onderzoek moesten kinderen steeds uit twee plaatjes het plaatje kiezen dat de betekenis van de zin het beste weergeeft. In het geval van (1) moesten kinderen dus kiezen uit een plaatje waarop een portier de deur opent en een plaatje waarop een postbode de deur opent. Serratrice vond dat tweetalige kinderen zinnen zoals in (1) vaker op de Engelse manier interpreteerden dan Italiaans eentalige kinderen. Zijn vond dus bewijs voor tussen-talige invloed.

Omdat er maar weinig onderzoeken bestaan naar taalbegrip, weten we niet hoe vaak de ene taal van tweetalige kinderen het begrijpen van de andere taal beïnvloedt, zoals Serratrice vond. Ook is het onduidelijk wat de rol is van taaloverlap, taaldominantie en leeftijd. Bovendien geven bestaande studies een incompleet beeld van de situatie. Dat komt omdat deze studies, op een uitzondering na (Lemmerth & Hopp, 2019), *offline* waren. Offline wil zeggen dat tussen-talige invloed alleen werd onderzocht in de *uiteindelijke* zinsinterpretatie van kinderen, zoals de keuze voor een plaatje. Zo'n soort offline taak heeft een aantal nadelen. Ten eerste laten offline taken ruimte voor expliciete kennis (Marinis, 2010). Doordat het kind alle tijd heeft om een beslissing te nemen, kan zij nadenken over een antwoord en misschien tot een andere interpretatie komen dan ze in eerste instantie had. Ten tweede laat een offline taak alleen het *eindproces* van zinsbegrip zien, nadat een hele zin is beluisterd (Marinis, 2010). Onderzoek laat echter zien dat we normaal niet wachten met het interpreteren van een zin tot we de hele zin hebben gehoord: we beginnen hier direct mee wanneer we het eerste stukje van een zin horen (Altmann & Steedman, 1988; Levelt, 1989). Het kan daarom goed zijn dat tussen-talige invloed al plaatsvindt *tijdens* het beluisteren van zinnen. Bestaande offline studies kunnen dit alleen niet laten zien, waardoor (subtiele) effecten van tussen-talige invloed mogelijk onzichtbaar blijven. Ten derde vragen offline taken vaak veel van het werkgeheugen van kinderen

(Marinis, 2010). Dit kan ervoor zorgen dat de uitkomst van zo'n taak niet altijd een goede weerspiegeling is van wat het kind daadwerkelijk kan.

Om goed te kunnen onderzoeken in hoeverre de ene taal van een kind het begrijpen van de andere taal beïnvloedt, zijn daarom zogenoemde *online* experimenten nodig. Zulke experimenten bouwen meer op impliciete processen, meten het zinsverwerkingsproces direct en vragen minder van het werkgeheugen (Marinis, 2010). Hierdoor zijn online experimenten geschikt om subtiele effecten van tussen-talige invloed *tijdens* zinsbegrip bloot te leggen. De hoofdvraag van dit proefschrift is daarom:

In hoeverre beïnvloedt de ene taal van een tweetalig kind het begrijpen van de andere taal tijdens de zinsverwerking?

Deze vraag hebben we onderzocht door middel van online experimenten in Hoofdstuk 3 en Hoofdstuk 5. Daarnaast onderzocht dit proefschrift vier deelvragen. Deze bespreek ik hieronder één voor één.

Deelvraag 1: In hoeverre wordt tussen-talige invloed in het online taalbegrip van tweetalige kinderen voorspelt door taaloverlap en taaldominantie?

Taaloverlap

In twee invloedrijk artikelen beargumenteerden Aafke Hulk en Natascha Müller (2000; Müller & Hulk, 2001) dat tussen-talige invloed alleen plaatsvindt wanneer er een bepaalde mate van overlap bestaat tussen de talen van tweetalige kinderen. In Taal A moet er sprake zijn van ambiguïteit: het lijkt voor het kind alsof er meerdere opties mogelijk zijn in de taal. In Taal B moet er juist geen sprake zijn van ambiguïteit: hier is maar één optie mogelijk. Daarbij moet de optie in Taal B overlappen met één van de opties in Taal A. Wanneer er sprake is van zo'n *gedeeltelijke overlap* tussen de talen, voorspellen Hulk en Müller invloed van Taal B op Taal A: kinderen zullen de gedeelde optie meer gaan gebruiken in Taal A. Een voorbeeld van gedeeltelijke overlap zijn bezittelijke constructies in het Frans en het Nederlands (gebaseerd op Nicoladis, 2012). Nederlands is hier Taal A. Zoals eerder uitgelegd, kan je in het Nederlands om uit te drukken dat iets van iemand is, bijvoorbeeld zeggen *de fiets van Els* of *Els haar fiets*. Frans is hier Taal B: alleen *le vélo de Els* ("de fiets van Els") is hier mogelijk. De voorspelling voor het Frans en Nederlands is dus dat tweetalige kinderen in het Nederlands vaker *de fiets van Els* zullen zeggen onder invloed van het Frans.

Tegelijkertijd zien we in bestaande studies dat tussen-talige invloed ook plaatsvindt zonder dat er *gedeeltelijke overlap* is (Foroodi-Nejad &

Paradis, 2009; Nicoladis, 2006). Het effect van taaloverlap op tussen-talige invloed is daardoor onduidelijk.

Taaldominantie

Tweetalige kinderen hebben typisch een dominante en een zwakkere taal (Grosjean, 1982). Verschillende onderzoeken laten zien dat de dominante taal de zwakkere taal beïnvloedt, maar niet andersom (Argyri & Sorace, 2007). Andere onderzoekers vonden juist wel tussen-talige invloed in beide richtingen. In sommige studies was de beïnvloeding groter van de dominante taal in de zwakkere taal dan andersom. In andere studies werd er daarentegen geen verschil tussen de dominante en zwakkere taal gevonden (Foroodi-Nejad & Paradis, 2009). Het precieze effect van taaldominantie op tussen-talige invloed is daarom door deze verschillende bevindingen ook onduidelijk.

Om een beter beeld te krijgen van het effect van taaloverlap en taaldominantie op tussen-talige invloed in bestaande studies, voerden we een meta-analyse uit in Hoofdstuk 2. De rol van taaloverlap in online zinsbegrip onderzochten we in Hoofdstuk 3. De rol van taaldominantie onderzochten we in Hoofdstuk 3 en 5.

Deelvraag 2: In hoeverre beïnvloedt het tweetalig-zijn het online zinsbegrip van tweetalige kinderen?

Onderzoek met volwassenen laat zien dat het online verwerken van een tweede taal minder efficiënt kan zijn dan een eerste taal (Clahsen & Felser, 2006; Hopp, 2010). Volwassen tweedetaalleerders doen er bijvoorbeeld soms wat langer over om zinnen te begrijpen dan moedertaalsprekers. Dit zou onder andere kunnen komen doordat mensen vaak minder ervaring hebben met hun tweede taal dan met hun moedertaal. Dit kan onder andere komen doordat ze minder taalaanbod in hun tweede taal krijgen en doordat ze minder vaardig zijn in hun tweede taal dan moedertaalsprekers (Sorace, 2011).

Sommige onderzoekers hebben gesuggereerd dat tweetalige kinderen ook minder efficiënt zijn in het verwerken van hun talen vergeleken met eentalige kinderen (Sorace & Serratrice, 2009). Het is echter onduidelijk of dit ook daadwerkelijk het geval is in het online taalbegrip. In dit proefschrift onderzochten we algemene effecten van tweetaligheid door altijd twee tweetalige groepen met verschillende taalcombinaties en een eentalige groep met elkaar te vergelijken. Als tweetalige kinderen minder efficiënt zijn tijdens de zinsverwerking verwachtten we dat de tweetalige groepen zich anders gedragen dan de eentalige groep: op dezelfde manier, ongeacht tussen-talige invloed. Dit onderzochten we in Hoofdstuk 3 en 5.

Deelvraag 3: In hoeverre is er tussen-talige invloed in het online zinsbegrip van tweetalige volwassenen?
Sommige onderzoeken laten zien dat tussen-talige invloed afneemt of zelfs verdwijnt in oudere kinderen (Argyri & Sorace, 2009; Serratrice, Sorace, Filiaci & Baldo, 2009; Unsworth, 2012). Ook suggereren sommige onderzoekers dat tussen-talige invloed een tijdelijke fase vormt in de taalontwikkeling (Hulk & Müller, 2000; Paradis & Genesee, 1996). Andere onderzoeken tonen daarentegen aan dat de talen van relatief oude kinderen (boven de acht jaar) elkaar kunnen beïnvloeden (Daskalaki, Chondrogianni, Blom, Argyri & Paradis, 2019). Bovendien vonden verschillende onderzoeken geen verband tussen tussen-talige invloed en leeftijd (Bosch & Unsworth, 2020; Nicoladis, 2002; 2003). Hierdoor is het onduidelijk in hoeverre tussen-talige invloed aanwezig blijft wanneer kinderen volwassen worden.

Er is tot nu toe weinig onderzoek gedaan naar tussen-talige invloed bij volwassenen die van jongs af aan met twee talen zijn opgegroeid (Kupisch, 2012; 2014). Daarbij meten weinig van deze onderzoeken online taalbegrip (maar zie Martohardjono, Phillips, Madsen II & Schwartz, 2018). Bovendien is er weinig bekend over de effecten van taaloverlap en -dominantie in de volwassen populatie (voor een uitzondering, zie werk van Tanja Kupisch: 2012; 2014). Daarom hebben we het online experiment in Hoofdstuk 3 herhaald in Hoofdstuk 4 met (jong)volwassen participanten.

Deelvraag 4: Hoe verhoudt tussen-talige invloed tijdens de zinsverwerking zich tot tussen-talige invloed in offline zinsbegrip?
Recente theorieën suggereren dat tussen-talige invloed in tweetalige kinderen ontstaat door co-activatie van talen *tijdens* de zinsverwerking (Nicoladis, 2006; 2012; Serratrice, 2016). Wanneer kinderen hun ene taal horen of spreken, is hun andere taal op de achtergrond ook actief (Koutamanis, Kootstra, Dijkstra & Unsworth, 2021; Von Holzen en Mani, 2012). Volgens Nicoladis (2006; 2012) zou dit kunnen verklaren waarom kinderen soms een woordvolgorde van de ene taal in de andere taal gebruiken. Doordat Taal A actief is tijdens het spreken van Taal B, kunnen kinderen soms onbedoeld een woordvolgorde uit Taal A kiezen. Zulke ideeën over co-activatie sluiten aan bij theorieën over tussen-talige invloed bij volwassen tweedetaalleerders (Hartsuiker & Bernolet, 2017).

Als tussen-talige invloed in zinsproductie en offline zinsbegrip inderdaad het gevolg is van co-activatie van talen tijdens de zinsverwerking, dan is de verwachting dat offline en online tussen-talige invloed hetzelfde patroon laat zien. Tegelijkertijd verwachtten we dat offline tussen-talige invloed alleen zichtbaar wordt, wanneer online co-activatie sterk genoeg is. Oftewel, we verwachtten dat tussen-talige invloed sterker is in online zinsbegrip dan in offline zinsbegrip. Dit hebben we getest in Hoofdstuk 5 door de uitkomsten van de online taak met de uitkomsten van een offline taak te vergelijken.

In de volgende secties bespreek ik de bevindingen van Hoofdstuk 2 tot en met 5. Daarna leg ik uit hoe we de resultaten van de verschillende hoofdstukken gecombineerd hebben in een model van tussen-talige invloed tijdens zinsverwerking (Hoofdstuk 6).

Hoofdstuk 2: De meta-analyse

In hoofdstuk 2 hebben we tussen-talige invloed in bestaande experimentele studies (productie en offline begrip) op een systematische manier onderzocht door middel van een meta-analyse. Daarbij hebben we getest in hoeverre taaloverlap, taaldominantie en leeftijd samenhangen met de sterkte van tussen-talige invloed. Hiervoor hebben we door middel van verschillende zoektermen meer dan 2000 mogelijk relevante studies geïdentificeerd. Van deze studies bleken 37 studies aan onze zoekcriteria te voldoen. Hiervan hadden we van 26 studies voldoende data om de meta-analyse uit te voeren. Deze 26 studies bevatten resultaten van in totaal 775 tweetalige kinderen (vergeleken met 750 eentalige kinderen), van 17 unieke taalcombinaties, 49 verschillende experimenten en 187 metingen van tussen-talige invloed. Voor iedere meting hebben we vervolgens een effectgrootte berekend (Hedges' *g*). Zo'n effectgrootte laat zien hoe sterk de talen van tweetalige kinderen elkaar beïnvloedden in een specifiek experiment. Hoe groter de effectgrootte, hoe sterker het effect.

Algemene analyse

In een eerste analyse hebben we de gemiddelde effectgrootte berekend van alle meting van tussen-talige invloed samen. Op deze manier onderzochten we in hoeverre tussen-talige invloed aanwezig was in de verschillende studies. We vonden een significante kleine tot gemiddelde effectgrootte van ongeveer 0.4 (*Hedges' g*). Dit resultaat laat zien dat tussen-talige invloed onderdeel is van tweetalig zijn. Tegelijkertijd liet onze analyse ook zien dat er grote verschillen waren tussen kinderen tussen en binnen dezelfde studies. We hebben vervolgens gekeken of we deze verschillen konden verklaren aan de hand taaloverlap, -dominantie en leeftijd.

Taaloverlap
Voor taaloverlap noteerden we voor iedere meting van tussen-talige invloed of het een situatie was van *gedeeltelijke overlap* of *geen overlap* tussen de talen van de kinderen. Het bleek dat verschillende studies verschillende definities voor overlap gebruikten. Om op een systematische manier overlap tussen studies te kunnen vergelijken, hebben we daarom een eigen definitie gehanteerd: we baseerden overlap op wat mogelijk is in de taal van volwassen moedertaalsprekers. Dit is relevant, omdat de oorspronkelijke definitie van Hulk en Müller (2000) juist uitgaat van het perspectief van het kind: in hoeverre staan kinderen meerdere opties toe in hun talen? Helaas hadden we te weinig informatie om overlap vanuit het kind te beoordelen.

We vonden geen significant effect van onze definitie van overlap: er was geen verschil tussen situaties van *gedeeltelijke* en *geen overlap*. Het is echter onduidelijk of de afwezigheid van een significant effect komt doordat taaloverlap geen rol speelt bij tussen-talige invloed of doordat taaloverlap niet vanuit het perspectief van het kind gedefinieerd was.

Taaldominantie
Voor taaldominantie noteerden we voor iedere meting van tussen-talige invloed of kinderen getest waren in hun dominante of in hun zwakkere taal. Ook hier bleek dat studies verschillende definities hanteerden. Om toch op een systematische manier naar dominantie te kunnen kijken, hebben we een objectieve manier gebruikt om dominantie te bepalen. Als kinderen getest werden in de taal van de maatschappij waarin ze woonden (bijvoorbeeld Nederlands in Nederland) dan noteerden we dat kinderen getest waren in hun dominante taal. Als kinderen in hun andere taal getest werden (bijvoorbeeld Turks in Nederland) dan rekenden we dat als hun zwakkere taal.

In de analyse vonden we een significant effect van taaldominantie: de dominante taal van kinderen beïnvloedde de zwakkere taal sterker dan andersom. Er is dus een relatie tussen de taaldominantie van het kind en tussen-talige invloed in bestaand onderzoek.

Leeftijd
Voor leeftijd hebben we per meting van tussen-talige invloed de gemiddelde leeftijd van de tweetalige kinderen genoteerd. Vervolgens hebben we getest of de talen van kinderen elkaar minder beïnvloeden naarmate kinderen ouder worden.

Er was een trend van leeftijd: hoe ouder kinderen waren, hoe minder de talen van kinderen elkaar beïnvloedden. Deze trend was echter niet significant. Dit resultaat suggereert dat tussen-talige invloed op oudere leeftijd niet verdwijnt.

Hoofdstuk 3: Cross-linguistic influence in het beluisteren van verschillende woordvolgordes – een online zinsverwerkingstaak

In dit hoofdstuk hebben we onderzocht hoe *snel* kinderen Nederlandse zinnen verwerken. We hebben twee type zinnen gebruikt: lange passieven (bijv., *de leeuw wordt geduwd door de beer*) en zinnen die beginnen met een bijwoordelijke bepaling (*Op de bank zingt de slang een liedje*). Lange-passiefzinnen in het Nederlands hebben een door-bepaling. Deze door-bepaling kan zowel *voor* (pre-werkwoord) als *na* het voltooid deelwoord komen (post-werkwoord; vergelijk de voorbeelden in 1 en 2).

(1) **pre-werkwoord**
De leeuw wordt **door de beer** geduwd en de muis rent door het bos.

(2) **post-werkwoord**
De leeuw wordt geduwd **door de beer** en de muis rent door het bos.

Volgens de Duitse grammatica is alleen de pre-werkwoordvolgorde correct (1), hoewel de post-werkwoordvolgorde soms ook in gesproken taal voorkomt. In het Engels is juist alleen de post-werkwoordvolgorde correct (2). De pre-werkwoordvolgorde overlapt dus *gedeeltelijk* tussen het Duits en het Nederlands en de post-werkwoordvolgorde overlapt *gedeeltelijk* tussen het Engels en het Nederlands.

In Nederlandse hoofdzinnen staat de persoonsvorm altijd op de tweede plaats. Dit is ook het geval wanneer een zin met een bijwoordelijke bepaling begint, zie het voorbeeld in (3). Een zin met de persoonsvorm op de derde plaats – na het onderwerp – is ongrammaticaal, zie het voorbeeld in (4). De plaatsing van de persoonsvorm overlapt *volledig* tussen het Nederlands en het Duits – Duits heeft dezelfde regel als het Nederlands. In het Engels, daarentegen, komt de persoonsvorm op een paar uitzonderingen na altijd *na* het onderwerp, zoals in (4). Er is dus *geen* overlap tussen het Nederlands en het Engels.

(3) Op de bank **zingt** de slang een liedje van school.
(4) *Op de bank de slang **zingt** een liedje van school.

De kinderen

In het experiment hebben we 40 Nederlands eentalige, 42 Duits-Nederlands en 39 Engels-Nederlands tweetalige kinderen zinnen laten beluisteren in het

Nederlands. Alle kinderen waren tussen de 5 en 9 jaar oud en de tweetalige kinderen waren begonnen met het leren van beide talen voor ze 4 jaar oud waren. Van alle tweetalige kinderen hebben we berekend hoe dominant ze waren in hun talen.

Het experiment

In dit hoofdstuk gebruikten we een *self-paced listening* experiment. In dit experiment kregen kinderen zinnen zoals in voorbeelden 1 tot en met 4 in stukjes te horen. Na een stukje gehoord te hebben, moesten kinderen op een knop drukken om het volgende stukje te beluisteren (zie voorbeeld 5, de schuine strepen geven de plekken weer waar de zin in stukjes is geknipt).

(5) De leeuw / wordt / door de beer/ geduwd / en / de muis / rent door het bos.

Op die manier konden we meten hoe lang kinderen erover deden om de verschillende stukjes van de zinnen te begrijpen. Het principe achter dit experiment is hoe langer het duurt om op de knop te drukken, hoe moeilijker het is om de zin te begrijpen.

De resultaten en conclusies

Algemeen

Een eerste bevinding was dat Duits-Nederlandse kinderen langzamer naar zinnen luisterden wanneer deze *gedeeltelijk* of *volledig* overlapten tussen hun talen. Dit suggereert dat het Duits het beluisteren van zinnen in het Nederlands beïnvloedt. We vonden geen bewijs voor tussen-talige invloed van het Engels in het Nederlands: de Engels-Nederlandse kinderen gedroegen zich hetzelfde als de eentalige kinderen.

Taaloverlap

Een tweede bevinding was dat taaloverlap uitmaakt voor tussen-talige invloed tijdens de zinsverwerking. De invloed van het Duits was namelijk het duidelijkst voor de *gedeeltelijk* overlappende lange-passiefzinnen. Invloed van het Duits op de *volledig* overlappende zinnen die met een doorbepaling begonnen was alleen zichtbaar wanneer we rekening hielden met de taaldominantie van de Duits-Nederlandse kinderen (zie volgende paragraaf). Er was geen tussen-talige invloed wanneer zinnen niet overlapten tussen het Duits en het Nederlands.

Taaldominantie
Een derde bevinding was dat taaldominantie een effect had op de invloed van het Duits op het Nederlands. Hoe sterker het Duits van een kind was ten opzichte van het Nederlands, hoe sterker de tussen-talige invloed was.

Effect van tweetaligheid
Als laatste zagen we geen negatief effect van tweetaligheid. In tegenstelling: de tweetalige kinderen beluisterden de zinnen juist sneller dan de eentalige kinderen. Als er dus al een effect van tweetaligheid was, was deze positief.

Hoofdstuk 4: Tussen-talige invloed in het beluisteren van verschillende woordvolgordes – volwassenen

In dit hoofdstuk hebben we het onderzoek in Hoofdstuk 3 herhaald met tweetalige (jong)volwassenen.

De deelnemers
Aan het experiment deden 25 eentalig opgegroeide Nederlandse en 25 Duits-Nederlands en 26 Engels-Nederlands tweetalige (jong)volwassenen tussen de 15 en 43 jaar mee. Onze verwachting was dat we nog steeds tussen-talige invloed zouden zien bij de oudere Duits-Nederlandse groep, maar mogelijk minder duidelijk dan bij de kinderen.

Het experiment
Het experiment dat we gebruikten in dit hoofdstuk was grotendeels hetzelfde als het experiment dat we met de kinderen hadden gebruikt. Het belangrijkste verschil met Hoofdstuk 3 was dat we in Hoofdstuk 4 een extra experiment hadden toegevoegd. In dit tweede experiment kregen de deelnemers niet alleen Nederlandse maar ook Duitse (voor de Duits-Nederlandse groep) of Engelse (voor de Engels-Nederlandse groep) zinnen te horen. Op deze manier wilden we de andere taal van de tweetalige groepen extra activeren.

De resultaten en conclusies

Algemeen
Een eerste bevinding was dat Duits-Nederlandse (jong)volwassenen zinnen trager beluisterden wanneer er *gedeeltelijke* overlap was tussen de talen. Deze vertraging was alleen zichtbaar in het experiment waarin de deelnemers ook Duitse zinnen hoorden. Dit resultaat suggereert aan de ene kant dat tussen-talige invloed tijdens de zinsverwerking blijft bestaan als kinderen opgroeien: we vonden hetzelfde effect bij (jong)volwassenen als we bij de

kinderen in Hoofdstuk 3 hadden gezien. Aan de andere kant laat het resultaat zien dat tussen-talige invloed subtieler is in (jong)volwassenen en alleen zichtbaar is wanneer de andere taal voldoende geactiveerd is.

Taaloverlap

Een tweede bevinding was dat taaloverlap uitmaakt: tussen-talige invloed was alleen zichtbaar in een situatie van *gedeeltelijke* overlap. In tegenstelling tot de resultaten van de kinderen in Hoofdstuk 3, vonden we geen bewijs voor invloed van het Duits op het verwerken van *compleet* overlappende zinnen in het Nederlands die met een bijwoordelijke bepaling beginnen.

Taaldominantie

Een derde bevinding was dat taaldominantie uitmaakt, net als bij de kinderen in Hoofdstuk 3: de invloed van het Duits op het Nederlands was alleen zichtbaar wanneer we rekening hielden met de taaldominantie van de (jong)volwassenen. Hoe sterker het Duits was ten opzichte van het Nederlands, hoe langzamer Duits-Nederlandse deelnemers lange-passiefzinnen (pre-werkwoord) beluisterden.

Effect van tweetaligheid

Een laatste bevinding was dat er geen duidelijk negatief effect was van tweetaligheid, net als bij de kinderen in Hoofdstuk 3. De tweetalige groepen – en met name de Engels-Nederlandse groep – waren over het algemeen iets langzamer in het beluisteren van zinnen dan de eentalig opgegroeide groep. Dit verschil tussen de groepen was echter niet significant.

Hoofdstuk 5: Tussen-talige invloed tijdens het interpreteren van persoonlijk voornaamwoorden – bewijs van oogbewegingen

In dit laatste empirische hoofdstuk hebben we onderzocht hoe Turks-Nederlands tweetalige kinderen persoonlijke voornaamwoorden in het Nederlands (*zij* en *hij*) interpreteren tijdens (online) en na (offline) het beluisteren van zinnen. In het Nederlands krijgen persoonlijk voornaamwoorden in zinnen zoals in (6) typisch een lokale interpretatie. Dat wil zeggen dat moedertaalsprekers van het Nederlands er vanuit zullen gaan dat Anna een slokje water neemt (Roberts, Gullberg & Indefrey, 2008).

(6) Anna en Sophie leren in de bibliotheek. Terwijl Anna een boek leest, neemt **zij** een slokje water.

In het Turks, daarentegen, zijn de voorkeuren anders (Azar & Özyürek, 2015). In de Turkse vertaling van zin (6) in (7) kan het voornaamwoord *o* niet aan Anna refereren. In plaats daarvan refereert het voornaamwoord in het Turks aan Sophie, de niet-lokale referent, of eventueel een niet-genoemde derde persoon. In Hoofdstuk 5 verwachtten we daarom dat Turks-Nederlands tweetalige kinderen meer geneigd zouden zijn om *zij* en *hij* in zinnen zoals (6) als niet-lokaal te interpreteren dan Nederlands eentalige kinderen, onder invloed van het Turks.

(7) Anna ve Sophie kütüphane çalışıyorlar.
Anna en Sophie bibliotheek werken
Anna kitap okurken, o sudan bir yudum alıyor.
Anna boek leest zij/hij water een slok neemt

De kinderen

Aan het onderzoek deden 17 Turks-Nederlands, 23 Duits-Nederlands tweetalige en 14 Nederlands eentalige kinderen mee. De kinderen waren tussen de 7 en 11 jaar oud en waren begonnen met het leren van beide talen voor ze 3 jaar oud waren. Net als in Hoofdstuk 3 en 4 hebben we berekend hoe dominant de kinderen waren in hun twee talen. Omdat Duits en Nederlands dezelfde voorkeuren hebben voor voornaamwoorden (Roberts et al., 2008) verwachtten we geen tussen-talige invloed in de Duits-Nederlandse groep kinderen.

Het experiment

Kinderen kregen verschillende varianten van het voorbeeld in (6) te horen. Bij ieder verhaaltje zagen ze drie plaatjes op een laptopscherm. In het geval van (6) een plaatje van Anna, van Sophie en van een boek. Op het moment dat kinderen het voornaamwoord in de zin hoorden, registreerde een eye-tracker (een soort camera) onderaan het laptopscherm waar kinderen naar keken. Op deze manier konden we erachter komen op welk moment kinderen naar welk plaatje keken tijdens het beluisteren van een zin. Het principe achter het eye-tracking experiment is dat wanneer kinderen naar zinnen luisteren, ze (onbewust) kijken naar wat ze horen. Hun ogen verraden dus hoe ze een zin interpreteren *tijdens* het beluisteren. Dus als een kind *zij* hoort en naar Anna kijkt, zouden we kunnen concluderen dat ze het voornaamwoord een lokale interpretatie geeft.

Aan het einde van ieder verhaaltje, stelden we kinderen een vraag. Bijvoorbeeld: “Wie neemt een slokje water?”. Door op een knop te drukken konden kinderen aangeven of dit de lokale persoon was (Anna) of de niet-lokale persoon (Sophie). Dit was onze offline maat.

De resultaten & conclusies

Algemeen

Een eerste bevinding was dat de Turks-Nederlandse kinderen als groep niet significant afweken van de andere twee groepen: ongeacht de voorkeuren in het Turks hadden de kinderen dezelfde kijkpatronen als de eentalige en de Duits-Nederlandse tweetalige kinderen. Maar als we taaldominantie meenamen in de analyses, zagen we wel bewijs voor tussen-talige invloed. Onder invloed van het Turks keken kinderen juist *minder* naar de niet-lokale persoon (Sophie). Deze bevinding was onverwacht, omdat de niet-lokale interpretatie de voorkeur heeft in het Turks. Wat dit resultaat laat zien is dat het Turks geactiveerd wordt tijdens het beluisteren van Nederlandse zinnen, maar dat deze activatie zo sterk onderdrukt wordt dat de optie die in het Turks de voorkeur heeft wegvalt.

Taaldominantie

Een tweede bevinding was dat de taaldominantie van kinderen uitmaakte: alleen wanneer taaldominantie werd meegenomen in de analyses zagen we effecten van het Turks op het Nederlands. Hoe sterker het Turks was ten opzichte van het Nederlands, hoe *minder* kinderen naar de niet-lokale persoon keken. Oftewel, hoe sterker het Turks, hoe minder kinderen de ‘Turkse strategie’ lieten zien.

Offline zinsbegrip

Een derde bevinding was dat de patronen die we in online zinsbegrip van kinderen zagen gereflecteerd werden in offline zinsbegrip. Met andere woorden, hoe sterker het Turks was ten opzichte van het Nederlands, hoe minder vaak kinderen de niet-lokale persoon kozen. Dit offline effect was echter niet significant, in tegenstelling tot het online effect.

Effect van tweetaligheid

Als laatste bevinding vonden we geen bewijs voor een algemeen effect van tweetaligheid. Dat wil zeggen dat we geen algemeen patroon zagen in de twee tweetalige groepen dat afweek van de eentalige groep, zowel online als offline.

Hoofdstuk 6 – Interpretatie van de resultaten aan de hand van ons zinsverwerkingsmodel (CLISP)

In Hoofdstuk 6 vatten we onze belangrijkste resultaten samen. Daarnaast stellen we in dit hoofdstuk een model voor om tussen-talige invloed in zinsverwerking te kunnen verklaren (CLISP). In deze sectie zal ik eerst kort per onderzoeksvraag onze conclusies geven. Vervolgens schets ik het CLISP model.

Hoofdvraag: tussen-talige invloed tijdens zinsverwerking in tweetalige kinderen

De experimenten in dit proefschrift laten bewijs zien voor tussen-talige invloed tijdens de zinsverwerking. Wanneer er een bepaalde mate van overlap is tussen de talen van tweetalige kinderen, resulteerde dit in Hoofdstuk 3 in langzamere zinsverwerking en in Hoofdstuk 5 in minder oogfixaties op de voorkeursinterpretatie van de andere taal. Tegelijkertijd laten deze resultaten zien dat tussen-talige invloed heel subtiel kan zijn en niet altijd met het blote oog zichtbaar is.

Deelvraag 1: effecten van taaloverlap en taaldominantie

Hoofdstuk 3 laat zien dat structuren die *gedeeltelijk* overlappen tussen talen mogelijk gevoeliger zijn voor tussen-talige invloed dan structuren die *compleet* overlappen. Voor structuren waarbij *geen* overlap was, vonden we geen bewijs voor tussen-talige invloed. Aan de ene kant sluiten onze bevindingen daardoor aan bij het voorstel van Hulk en Müller (2000) over taaloverlap. Aan de andere kant laten onze resultaten zien dat *gedeeltelijke* overlap niet *noodzakelijk* is voor beïnvloeding tussen talen.

Hoofdstuk 3 en 5 suggereren verder dat taaldominantie een belangrijke rol speelt voor het plaatsvinden van tussen-talige invloed tijdens de zinsverwerking. Ten eerste vonden we in beide studies sterkere effecten van tussen-talige invloed naarmate kinderen dominanter waren in de andere taal. Ten tweede was de invloed van het Duits en Turks op het Nederlands (soms) alleen zichtbaar wanneer we naar taaldominantie keken. Onze bevindingen over taaldominantie passen bij de resultaten van onze meta-analyse en bij resultaten van individuele productie en offline studies (Serratrice, 2013).

Deelvraag 2: algemeen effect van tweetaligheid

Hoofdstuk 3 en 5 laten geen bewijs zien voor een algemeen negatief effect van tweetaligheid. Dit is in tegenstelling tot online studies met volwassen tweedetaalleerders (Hopp, 2010) en tot wat in eerdere studies gesuggereerd is voor tweetalige kinderen (Sorace & Serratrice, 2009). Het enige algemene effect dat in Hoofdstuk 3 naar voren komt is dat tweetalige kinderen juist sneller zinnen verwerken dan eentalige kinderen. Er moet wel een kanttekening worden geplaatst bij onze bevindingen: alle kinderen die meededen aan de experimenten, woonden in en gingen, op één kind na, naar school in Nederland. Hierdoor was het taalniveau en de hoeveelheid aanbod in het Nederlands van deze kinderen relatief hoog. Mogelijk zouden kinderen die niet in Nederland wonen en naar school gaan, minder efficiënt kunnen zijn in het verwerken van Nederlandse zinnen dan eentalige kinderen.

Deelvraag 3: tussen-talige invloed in tweetalige volwassenen
Hoofdstuk 4 laat zien dat tussen-talige invloed tijdens de zinsverwerking aanwezig blijft in (jong)volwassen tweetaligen. Dit komt overeen met onze bevindingen over leeftijd in de meta-analyse en met bestaande studies die tussen-talige invloed in tweetalige volwassenen vonden (Kupisch, 2012; Martohardjono et al., 2017). Ook suggereren onze resultaten dat taaloverlap en taaldominantie bij volwassenen een rol blijven spelen. Wel lijkt tussen-talige invloed subtieler te zijn in volwassenen en alleen tot uiting te komen wanneer de andere taal voldoende is geactiveerd.

Deelvraag 4: tussen-talige invloed in offline zinsbegrip
Hoofdstuk 5 laat zien dat tussen-talige invloed in offline zinsbegrip hetzelfde patroon volgt als tussen-talige invloed in online zinsbegrip. Tegelijkertijd is het offline effect niet significant, terwijl het online effect dat wel is. Deze resultaten sluit aan bij onze verwachting dat offline tussen-talige invloed een zwakkere reflectie is van online tussen-talige invloed.

Het CLISP model
Voor zover wij weten, kunnen bestaande modellen van tussen-talige invloed de resultaten in ons proefschrift niet verklaren (Hartsuiker & Bernolet, 2017; Nicoladis, 2006; 2012; Serratrice, 2016). Daarom stellen wij in Hoofdstuk 6 een eigen model voor, gebaseerd op verschillende bevindingen van eerder onderzoek naar online zinsverwerking (CLISP).

Het model bestaat uit vier componenten. De eerste component beschrijft de manier waarop de talen van tweetalige kinderen schematisch gerepresenteerd zijn in het brein. Onder andere op basis van modellen van Hartsuiker en collega's (Hartsuiker, Pickering & Veltkamp, 2004; Hartsuiker & Bernolet, 2017) gaan wij er vanuit dat woordrepresentaties (lemma's) in het lexicon verbonden zijn met onder andere de betekenissen (semantische representaties) en mogelijke abstracte woordcombinaties waarin de lemma's voor kunnen komen. We verwachten bijvoorbeeld dat het Nederlandse lemma voor *zingen* in het hoofd van een kind verbonden is met de betekenis voor zingen en met de woordvolgordes waarin het woord voor kan komen. Daarnaast nemen we aan, op basis van studies met volwassenen en kinderen (Hartsuiker et al., 2004; Hsin, Legendre & Omaki, 2013; Vasilyeva et al., 2010), dat wanneer twee talen identieke woordvolgordes hebben, de representaties hiervoor in het hoofd van een kind gedeeld worden (zoals pre-werkwoord passieven in het Duits en het Nederlands). Bovendien gaan we er vanuit dat de betekenissen van woorden ook gedeeld worden tussen talen (zoals de betekenis voor *zingen* in het Duits en het Nederlands).

De tweede component van het model beschrijft de activatie van talen tijdens het zinsverwerkingsproces. Hierbij is de eerste aanname van het model dat wanneer kinderen zinnen horen, de bijbehorende lemma's actief worden. Vervolgens stroomt deze activatie naar de mogelijke betekenissen van deze lemma's en de abstracte woordcombinaties (Dell, 1986; Pickering & Branigan, 1998). Wanneer kinderen bijvoorbeeld een lange-passiefzin horen met een pre-werkwoord door-bepaling (*de leeuw wordt door de beer geduwd*), worden dus alle lemma's van de individuele woorden actief, de betekenissen van de individuele woorden en de abstracte pre-werkwoord woordvolgorde. Vervolgens is de tweede aanname van het model dat lemma's van de andere taal ook actief worden. Dit komt omdat de activatie van gedeelde betekenissen en abstracte woordvolgordes doorstroomt naar lemma's ongeacht tot welke taal deze lemma's behoren (Hartsuiker et al., 2004; Hartsuiker & Bernolet, 2017). Met andere woorden, elementen van beide talen worden actief in tweetalige zinsverwerking.

De derde component van het model beschrijft processen die nodig zijn om de activatie van beide talen te controleren. Het proces waar wij in ons model op focussen is *inhibitory control*: het blokkeren van activatie (Green, 1998; Miyake et al., 2000). CLISP gaat er vanuit dat tweetalige kinderen en volwassenen de activatie van de andere taal zullen proberen tegen te gaan. Dit doen ze door activatie van lemma's te *inhiberen* (Green, 1998; Green & Abutalebi, 2013). Dit verklaart onze online resultaten in Hoofdstuk 3, 4 en 5: het blokkeren van activatie van lemma's van het Duits en Turks heeft als gevolg dat betekenissen en abstracte woordvolgordes die verbonden zijn met deze lemma's ook (deels) geblokkeerd worden. Daardoor kunnen tweetaligen deze representaties minder makkelijk selecteren. Dit leidt dan tot vertragingen tijdens de zinsverwerking (Hoofdstuk 3 en 4) en *minder* oogfixaties op de interpretatie die bij het Turks past (Hoofdstuk 5).

De vierde en laatste component van CLISP beschrijft verwerkingscapaciteit. Om zinnen te kunnen begrijpen en produceren is het werkgeheugen betrokken (Just & Carpenter, 1992). Hoe meer verschillende lemma's en bijbehorende representaties geactiveerd – en geblokkeerd – worden, hoe meer het werkgeheugen belast wordt (Conway & Engle, 1994; Just & Carpenter, 1992). Dit kan er aan de ene kant voor zorgen dat kinderen *online* minder capaciteit overhouden voor het interpreteren van zinnen. Dit kan tot vertragingen leiden (zoals gebeurde in Hoofdstuk 3). Aan de andere kant kan een sterke belasting van geheugencapaciteit erin resulteren dat kinderen soms offline een abstracte woordvolgorde of een betekenis van de andere taal kiezen. Dit laatste verklaart bevindingen van tussen-talige invloed in zinsproductie en offline -begrip.

Het CLISP-model verklaart de verschillende effecten van tussen-talige invloed die we in Hoofdstuk 2 tot en met 5 zien. Daarnaast verklaart CLISP de rol van taaloverlap, taaldominantie en leeftijd, en de verschillen die we zien tussen online en offline tussen-talige invloed. Bovendien maakt CLISP concrete voorspellingen over tussen-talige invloed die in toekomstige studies getest kunnen worden.

Conclusies

In dit proefschrift hebben we laten zien dat de talen van tweetalige kinderen elkaar beïnvloeden. Eerder onderzoek toont dat tussen-talige invloed soms duidelijk zichtbaar is doordat een tweetalig kind iets op een andere manier zegt dan een eentalig kind. Dit proefschrift laat zien dat beïnvloeding tussen talen ook heel subtiel kan zijn en soms alleen zichtbaar wordt wanneer er wordt gekeken wat er *tijdens* taalverwerking gebeurt. Tegelijkertijd suggereren onze resultaten dat tweetalige kinderen hun talen goed uit elkaar kunnen houden: aan de ene kant zijn de talen van tweetalige kinderen telkens in samenspel, maar aan de andere kant lijken kinderen deze co-activatie snel en efficiënt te kunnen sturen en inhiberen.

De resultaten van dit proefschrift laten ook zien dat niet ieder kind even gevoelig is voor tussen-talige invloed tijdens zinsverwerking. Hoe dominanter een taal van een kind, hoe groter de kans dat deze de verwerking van de andere taal (zichtbaar) beïnvloedt. Bovendien lijkt er een bepaalde mate van overlap tussen talen nodig te zijn om elkaar online te beïnvloeden. Dit zou kunnen betekenen dat talen die (syntactisch en/of morfologisch) meer op elkaar lijken, elkaar sterker beïnvloeden dan talen die veel van elkaar verschillen. Als laatste laat het proefschrift ook zien dat tussen-talige invloed blijft bestaan in (jong)volwassenen. Het is dus niet zo dat kinderen 'over tussen-talige invloed heen groeien'. In plaats daarvan lijken de talen van (jong)volwassenen elkaar op een subtielere manier te beïnvloeden.

In het kort toont dit proefschrift aan dat tussen-talige invloed er nou eenmaal bij hoort wanneer een kind tweetalig opgroeit. Talen beïnvloeden elkaar mogelijk zelfs meer dan we ons eigenlijk realiseren. Tegelijkertijd is deze beïnvloeding vaak heel subtiel en onzichtbaar. Deze bevindingen laten zien dat tweetalige kinderen heel efficiënt met hun talen kunnen omgaan en uitgroeien tot zeer getrainde taalgebruikers.

Dankwoord

Eind november 2015 schreef Sharon mij een korte mail, waarin ze vroeg of ik misschien een onderzoeksvoorstel in wilde dienen voor Promoties in de Geesteswetenschappen. Een aantal eerdere onderzoeksvoorstellen voor andere beurzen waren in de afgelopen jaren afgewezen, mijn contract als onderzoeksassistent in Utrecht was afgelopen en ik stond op het punt om op reis te gaan naar Zuid-Amerika. Maar goed, waarom niet nog een (laatste) poging wagen? Vlak voor ik op het vliegtuig stapte, dienden we mijn voorstel in. Mijn gesprekken met CLS voerde ik om 5 uur 's ochtends, met Sharon en Ton overlegde ik vanuit mijn hostel in Colombia en de pitch voor NWO oefende ik tijdens mijn hikes door de Andes in Ecuador. Blijkbaar had mijn reis een positieve uitwerking op mijn aanvraag: ik kreeg de beurs.

Zo soepel als de aanvraag van mijn project verliep, zo wisselvallig bleken de vier en een half jaar van mijn promotieonderzoek te zijn. Sommige perioden kwam ik goed vooruit en kreeg ik positieve reacties op mijn werk. Op andere momenten liep niet alles even soepel. Het vinden van voldoende tweetalige kinderen bleek soms lastig, mijn planning was standaard te optimistisch, waardoor ik vaak het gevoel had achter de feiten aan te lopen en corona gooide roet in het eten wat betreft mijn laatste dataverzameling. Ondanks alles is het goed gekomen. En dat heb ik voor een heel groot deel aan alle steun van de mensen om mij heen te danken.

Allereerst natuurlijk Sharon en Ton. Sharon, bedankt voor al je hulp, je vertrouwen in mij en je geduld bij het lezen en becommentariëren van alle vijftien versies(!) van mijn eerste paper. Zonder jou was ik nooit aan dit project begonnen en had ik het ook nooit af gekregen. Een betere supervisor had ik me niet kunnen wensen! Ton, voor ik jou kende, dacht ik dat promotoren eens in de paar maanden om de hoek komen kijken en zich verder niet met hun promovenda bemoeien. Je hebt me positief verrast met je betrokkenheid. Ik vond het ontzettend fijn om met je samen te werken en heb heel veel van je geleerd. Bedankt voor al je hulp en ook je gezelligheid!

Natuurlijk wil ik ook de (andere) leden van het 2in1 project bedanken. Elly, ik vond het ontzettend leuk om gezamenlijk aan onze projecten te werken. Ik vind ik het heel fijn dat je mijn (virtuele) kamergenoot en paranimf bent. Bedankt dat ik altijd bij jou mijn hart kon luchten de afgelopen jaren. Figen, I'm so happy that you joined our project and that you are my paranimf. It was great to already meet you beforehand in Edmonton. I love our discussions – often in the virtual kitchen – about research, cows and all topics in between. Thank you for making me think about the results of this thesis and

CLISP in different ways. Gerrit Jan, bedankt voor alle gezelligheid, je optimistische houding en de grapjes. Heb je nog steeds een laadje vol drop in je nieuwe kantoor? Christa, Judith, Lisa en Kyra, bedankt voor het inplannen van alle testafspraken en het regelen en organiseren van allerlei andere belangrijke dingen. Zonder jullie was het nooit gelukt om zoveel kinderen en volwassenen te vinden en te testen. Natuurlijk ook bedankt aan alle testers en families die hebben meegewerkt aan mijn onderzoek.

Ook wil ik al mijn collega-PhDs bedanken. Allereerst de oorspronkelijke 'bewoners' van mijn buurkantoor: Claire, Ferdy en Jan Willem. Claire en Jan Willem, het was en is ontzettend fijn om vanaf het begin lotgenoten te hebben, die weten hoe 'speciaal' het is om onderzoek te doen met kinderen. Ik kijk met veel plezier terug op al onze gesprekken. Ferdy, bedankt voor alle "koffie?!" wanneer ik even helemaal klaar was met alles. Saskia, als nieuwe kantoorgenoot doe je zeker niet onder voor de andere drie. Bedankt voor alle koffiepauzes, offline en online, samen sporten en je verhalen, vooral over de duiven. Hanno, ook dankjewel voor alle (online) gezelligheid, samen boulderen en natuurlijk Gathertown! Katherine, Aurora, Lisa and Chen, our morning and afternoon Zoom meetings were a great help during the last phase of my thesis. Thanks for your positivity and all the fun conversations. Katherine, extra credits go to you of course for having been there for me almost every day! Theresa, bedankt voor de wandelingen, onze gesprekken over *Maultaschen* en 'het brengen van contrasten' in mijn leven. Wei, thanks for having (online) coffee breaks with me, discussing all different kinds of Chinese food, and for inviting me to the hotpot dinner at your place. Emily, Martijn, Lotte, Tim, Yu en Xiaoru, bedankt voor alle gesprekken in de keuken over van alles en nog wat.

In dit dankwoord mag natuurlijk mijn illustere kamergenoot niet ontbreken. Louis, je hebt een extra dimensie toegevoegd aan mijn tijd in het Erasmusgebouw. Ik kijk met ontzettend veel plezier terug op mijn jaren in ons kantoor: door de vlaggetjes voor mijn verjaardag, onze gesprekken over chocola en alles, jouw papierverzameling die een eigen leven leidt. Ik steel je quote: "Blijf zoals je bent".

Aan iedereen van de CDAM onderzoeksgroep, bedankt voor de inspiratie en al jullie feedback op mijn onderzoek de afgelopen jaren. Susanne, Anna and Elise, thank you for being involved in some of the studies in this thesis. It was great to work with you. Margret, Bob en Henk, bedankt voor de technische ondersteuning.

Natuurlijk was mijn ervaring de afgelopen jaren in het Erasmusgebouw niet hetzelfde geweest zonder alle koffiepauzes. Ik ga hiervoor geen namen noemen, want dan ben ik bang dat ik sommigen per

ongeluk vergeet. Aan iedereen die regelmatig om 11:00 in de keuken op de 8e verdieping was, bedankt. Dit is één van de onderdelen van mijn PhD waar ik met heel veel plezier aan terugdenk. Gelukkig konden we de koffiepauzes het afgelopen jaar online voortzetten. Wessel, dankjewel dat je standaard op donderdag en vrijdag mijn mede-talkshow-host bent.

Ook een dankjewel aan degenen buiten de RU, die ervoor zorgden dat ik me ook met hele andere dingen bezig moest houden dan onderzoek. Ik wil hier een paar mensen extra bedanken, maar natuurlijk is de lijst veel langer. Yuri, Raëma, Erik, Egmond, bedankt voor al het rennen, survivallen en boulderen. Sara, de beste huisgenoot die er is, bedankt voor alle gezellige etentjes en avondjes “Ik vertrek” kijken. Anne, bedankt voor alle high-teas, we moeten snel weer gaan! Jack, bedankt dat je altijd zin hebt om iets samen met mij te ondernemen. Daisy, bedankt voor je mooie tekeningen voor mijn eerste experiment. Frank, bedankt dat je om de zoveel tijd even checkte of het nog allemaal goed ging met me. Eva, danke, dass du dich so gut um Inshirah gekümmert hast als ich in Nijmegen war. En niet alleen een bedankt, maar ook een sorry voor iedereen die nog steeds niet weet hoe het nou zit met die ganzen, eenden en krokodillen...

Pap en mam, jullie hebben me altijd gesteund. Bedankt daarvoor! Ik vond het ontzettend fijn dat jullie naar Nijmegen kwamen wanneer ik het te druk had om naar jullie te gaan. Milan, Axel en Sasja, bedankt voor alle zin en onzin in de app. Oma, dankjewel voor de gezellige avondjes in Dieren. Susanne, Andreas, Sarah, Peter und Inge, danke, dass ihr mir helft, mich in Deutschland zu Hause zu fühlen.

Felix, you were there for me every step of the way. Even me being at home for a year working (more than) full-time on my thesis did not scare you off. Thank you for everything!

Curriculum Vitae

Chantal van Dijk was born on September 26, 1988 in Apeldoorn, the Netherlands. She obtained her bachelor's degree in Liberal Arts and Sciences and her master's degree in Linguistics (cum laude) from Utrecht University. During her master's she spent a semester abroad at Bangor University and she also worked as a research assistant on a project investigating bilingual children at Utrecht University. After her studies she worked as a research assistant at Utrecht University on a project on WhatsApp language. In 2016 she received a grant from NWO (Promoties in de Geesteswetenschappen) to conduct her PhD project starting the same year at the Centre for Language Studies at the Radboud University in Nijmegen. She is currently employed as a post-doctoral researcher and lecturer at the same university.