Speech production, perception, and input of simultaneous bilingual preschoolers:

Evidence from voice onset time

Antje Stoehr

ISBN: 978-90-76203-90-4

Cover: Gesa Braun
Printing: Ipskamp Printing, Enschede

Speech production, perception, and input of simultaneous bilingual preschoolers:

Evidence from voice onset time

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 donderdag 1 maart 2018
om 12:30 uur precies

door

Antje Stoehr
geboren op 25 maart 1987
te Bad Oldesloe (Duitsland)

Promotoren

Prof. dr. J.P.M. Fikkert
Prof. dr. J.G. van Hell (The Pennsylvania State University, Verenigde Staten)

Copromotor

Dr. A.T. Benders (Macquarie University, Australië)

Manuscriptcommissie

Prof. dr. J.J.M. Schoonen (voorzitter)
Prof. dr. A. De Houwer (Universität Erfurt, Duitsland)
Dr. M. Kehoe (Université de Genève, Zwitserland)
Prof. dr. W.M. Lowie (Rijksuniversiteit Groningen)
Dr. E. Simon (Universiteit Gent, België)

Speech production, perception, and input of simultaneous bilingual preschoolers:

Evidence from voice onset time

Doctoral Thesis

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. J.H.J.M. van Krieken,
according to the decision of the Council of Deans
to be defended in public on Thursday, March 1, 2018
at 12:30 hours

by

Antje Stoehr
Born on March 25, 1987
in Bad Oldesloe (Germany)

Supervisors
Prof. Dr. J.P.M. Fikkert
Prof. Dr. J.G. van Hell (The Pennsylvania State University, USA)

Co-supervisor
Dr. A.T. Benders (Macquarie University, Australia)

Doctoral Thesis Committee
Prof. Dr. J.J.M. Schoonen (chair)
Prof. Dr. A. De Houwer (Universität Erfurt, Germany)
Dr. M. Kehoe (Université de Genève, Switzerland)
Prof. Dr. W.M. Lowie (Rijksuniversiteit Groningen)
Dr. E. Simon (Universiteit Gent, Belgium)

Table of contents

Chapter 1

General introduction

Bilingual first language acquisition is a process in which children acquire two languages from birth (De Houwer, 1990; Meisel, 1989). In this simultaneous bilingual acquisition process, children acquire two native languages in parallel, rather than a first language (L1) and a second language (L2) sequentially.

The field of bilingual first language acquisition emerged in the early twentieth century through the seminal work of Ronjat (1913) and Leopold (1939-1949), who reported on the longitudinal language development of a French-German bilingual child and an English-German bilingual child. Ronjat's son Louis and Leopold's daughter Hildegard acquired two languages from birth because their parents spoke different native languages. Both children were brought up in a linguistically similar fashion, namely using the one parent-one language strategy, in which each parent consistently speaks only their native language to the child (Ronjat, 1913). Yet, the children's acquisition processes seemed to differ: Louis acquired French and German fast and with few signs of confusion, while Hildegard passed through a stage in which she seemed to mix English and German. While the study of bilingual first language acquisition was still in its infancy, it revealed remarkable individual differences within the acquisition processes of bilingual children.

Language exposure is one factor that can account for differences between the acquisition processes of bilingual children (De Houwer, 2011; Gathercole, Thomas, Roberts, Hughes & Hughes, 2013). In terms of language exposure, two main settings of simultaneous bilingual acquisition must be distinguished: bilingual acquisition in a bilingual sociocultural setting and bilingual acquisition in a monolingual sociocultural setting. In some bilingual environments, such as Barcelona, Brussels, Luxembourg and Montreal, bilingual children are exposed to two languages in their broader social environment as well as possibly in their homes. In such acquisition contexts, bilinguals may develop a relatively balanced type of bilingualism with very high proficiency in both languages.

By contrast, a child's bilingualism can also emerge in a monolingual environment when the parents speak different native languages, as was the case for Hildegard Leopold and Louis Ronjat. When a bilingual child grows up in an environment in which only one language is spoken in the community, the exposure to the language that is not spoken in the community might only be provided by one parent. In such a scenario, the exposure to both languages will unlikely be equal because the caregivers typically do not spend an equal amount of time with the child, and later on, only one language is spoken at daycare, school, and the broader social environment. When only one language is spoken in the community, bilinguals often have imbalanced exposure to their two languages. This imbalance makes it unlikely for a bilingual child to develop equal proficiency in both languages when raised in a non-bilingual society (De Houwer, 2009). To capture the differences between simultaneous bilinguals raised in a bilingual environment and those simultaneous bilinguals who acquire a minority language in a majority language context, the term *heritage speaker* has been introduced (Cummins, 2005; Kupisch, 2013; Kupisch & Rothman, 2016; see Montrul, 2016, p. 40, for an overview). In this dissertation, the focus will be on children aged between 3;6 and 6;0 (years; months) who grow up as heritage speakers of German in the Netherlands.

Heritage speakers are generally dominant in the language of the environment and their proficiency in the heritage language can vary largely between individuals (Valdés, 2000a, 2000b). While some heritage speakers only understand the heritage language without speaking it (e.g., Au, Knightly, Jun & Oh, 2002), others are relatively balanced bilinguals with similar proficiency in both languages (e.g., Kupisch, Akpinar & Stöhr, 2013; Kupisch, Lein, Barton, Schröder, Stangen & Stoehr, 2014). Hildegard Leopold and Louis Ronjat acquired German as a heritage language in the USA and France, respectively. Differences in their heritage language exposure may have been related to the observed differences between their language acquisition trajectories. Hildegard was exposed to German from her father and only visited Germany twice before she entered school. Louis' exposure to German came from more sources, including his German mother and the family's servants from Germany, and Louis also visited his extended German family frequently. In sum, individual differences in acquisition outcomes in the heritage language provide a test case for the effects of the amount of language exposure in simultaneous bilingual acquisition.

Decades after the first reports on the language acquisition process of bilingual children, researchers rediscovered the field of bilingual first language acquisition (e.g., Arnberg, 1981; De Houwer, 1990; Genesee, 1989; Meisel, 1989; Volterra & Taeschner, 1978), possibly inspired by the increasing interest in (monolingual) language acquisition research and the discipline of psycholinguistics in general (Lambert, 1981; see Levelt, 2014, for an overview). In contrast to Ronjat's and Leopold's diary studies, this more recent line of bilingualism research benefitted from technological advancements. These allowed for studying bilingual children's language production using experimental methods and acoustic analyses (see Genesee & Nicoladis, 2005, for an overview of studies). Towards the end of the twentieth century, researchers started to investigate bilingual infants' speech perception, and this line of research provided first evidence for early language differentiation in perception (Bosch & Sebastián-Gallés, 1997). Yet, relatively little is known about simultaneous bilingual children's language acquisition to this day, and there are three primary reasons which complicate drawing generalizable conclusions across existing studies (see Kehoe, 2015, for a review). First, the field is predominantly based on case studies or studies using small heterogeneous samples. Second, many studies lack monolingual control groups for each language the bilinguals speak. These two reasons make it difficult to draw generalizable conclusions of the effect of bilingualism on language development independent from individual differences among children that affect developmental trajectories, be they bilingual or monolingual. Third, the vast majority of bilingualism studies has been conducted on bilingual children who were heritage speakers of Spanish and acquired English as the majority language. It is therefore difficult to determine to what extent previous findings can be generalized to bilingual acquisition of different language combinations.

Even though the field of bilingual first language acquisition is restricted in its generalizability, important conclusions can be drawn from previous work. Most researchers agree that bilingual children pass the same linguistic milestones within a similar age range as monolingual children, including the onset of babbling, production of first words, and rate of vocabulary growth (e.g., De Houwer, Bornstein & Putnick, 2014; Nicoladis & Genesee, 1997; Oller, Eilers, Urbano & Cobo-Lewis, 1997; Pearson, Fernández, Lewedag & Oller, 1997; Werker & Byers-Heinlein, 2008).

Nevertheless, one common observation is that phonetic aspects in the speech of bilingual children seem to diverge from monolingual children's speech in at least one language, which reflects the well-known notion that bilinguals are not two monolinguals in one person (Grosjean, 1989). What is the origin of these observed differences between the speech of bilinguals and monolinguals? The initial assumption that bilingual children start out speaking one mixed language (Volterra & Taeschner, 1978) has largely been replaced by the view that bilingual children have two linguistic systems that may interact. These linguistic systems are thus non-autonomous, which can lead to acquisition outcomes that differ from those of monolingual children (Paradis & Genesee, 1996). While the concept of cross-linguistic influence is widely accepted (Fabiano & Goldstein, 2005; Fabiano-Smith & Bunta, 2012; Kehoe, 2002; Kehoe, Lleó & Rakow, 2004; Kellerman & Sharwood Smith, 1986; Lleó & Kehoe, 2002; Müller & Hulk, 2001; Paradis & Genesee, 1996), it is presumably not the sole explanation for bilingual children's deviation from their monolingual peers (Genesee & Nicoladis, 2005; Werker & Byers-Heinlein, 2008). Differences in the phonetic characteristics in the speech of bilingual and monolingual children may also be related to differences in the amount of language input per language, as well as the difference between monolingual children's input from native speakers versus bilingual children's (partially) non-native input. The influence of these quantitative and qualitative aspects of the input on bilingual children's phonological acquisition is addressed in this dissertation.

The research conducted for this dissertation is situated at the interface between phonetics and phonology, and focuses on the acquisition of the voicing contrast of Dutch-German bilingual children through measures of voice onset time (VOT). VOT is the primary phonetic cue to distinguish 'voiced' plosives (such as /b/) from 'voiceless' plosives (such as /p/). Throughout this dissertation, the notations 'voiced' and 'voiceless' with single quotation marks will be used to refer to the phonological status of plosives. Both Dutch and German have a phonological voicing contrast, but they differ in its phonetic implementation, which allows for measuring potential cross-linguistic effects in bilingual children's speech (see Chapter 2, section 1.1 for a detailed description of VOT and voicing contrasts). In this dissertation, I take a multi-method approach and use evidence from children's speech production, speech perception, and parental language input. The combination of these measures allows for determining whether simultaneous bilingual children speaking Dutch and German show evidence for language-specific and monolingual-like phonological systems in both languages in *production* and whether the *amount of language exposure* and *non-native input* are associated with bilingual children's VOT production. In addition, the children's *perception* of the voicing contrast was tested to disentangle production and perception effects, and to determine if bilingual children associate VOT with voicing categories differently in Dutch and German.

Bilingual preschoolers were chosen as a target population for two primary reasons. First, much of the research on simultaneous bilingual phonological acquisition has either been conducted with infants and toddlers, which is likely related to the observation that about one quarter of simultaneous bilingual children stops speaking the heritage language after they start attending (pre)school (De Houwer, 2007, 2009). Those bilingual children who keep speaking two languages provide novel and important information on bilingual development over time.

Second, it is assumed that phonological representations are built in the mental lexicon in the course of word learning (Fikkert, 2010; Werker & Curtin, 2005), which makes it important to study bilingual children past the initial vocabulary spurt because phonological representations may get refined in the course of lexical development.

Before addressing speech production, speech perception and the input of bilingual children in more detail, the following section defines core terminology that is used throughout this dissertation. In the final section of this chapter, I provide the specific research questions and the outline of this dissertation.

1 Terminology used in this dissertation

A recurring issue in the bilingualism literature is the use of ambiguous terminology (Kupisch & Rothman, 2016; Rothman & Treffers-Daller, 2014). This section is therefore dedicated to defining seven key terms used in this dissertation.

Language-specific: This dissertation is concerned with phonemes that are shared between Dutch and German, such as /p/. Yet, /p/ is phonetically implemented with different VOT values in Dutch and German. If a bilingual produces [p] differently in Dutch than in German, these productions are considered language-specific, irrespective of whether the exact VOT values differ from monolingual productions.

Monolingual-like: This term is used when the acquisition process or outcome of bilinguals is compared to that of monolinguals. When I refer to bilingual children as being (not) monolingual-like, this is in reference to age-matched monolingual children.

Native-like versus non-native-like: An L2 learner can be described as native-like when speaking the L2 similar to someone who speaks this language as L1. In this dissertation, I keep native-likeness distinct from monolingual-likeness. To determine whether a German mother of a bilingual child achieved native-like VOT in Dutch production, she can be compared to a native speaker of Dutch, who speaks German as L2. If differences are observed, it can be concluded that her VOT productions in Dutch are non-native-like. To evaluate L2 attainment, bilingual native-speakers appear to be a more suitable reference point than monolinguals, because they presumably accommodate a similar number of phonemes as the L2 speaker.

Differential acquisition: If bilingual children or adults produce VOT in a non-monolingual-like fashion, their productions can be referred to as differential

(Kupisch & Rothman, 2016). As the term indicates, differential acquisition neutrally refers to a difference rather than implying any kind of delay or incompleteness.

Target-like versus non-target-like: Target-likeness describes the speech of a child or language learner in reference to the input that is available to them. For example, the German language input that is available to a Dutch-German bilingual child who grows up in the Netherlands is most likely different from the input of a monolingual child raised in Germany.

Input quantity: The input quantity refers to a bilingual's proportional amount of exposure to each language at the time of testing, and is measured by the Bilingual Language Experience Calculator (BILEC; Unsworth, 2013) in this dissertation. The input quantity of both languages always adds up to 100%, for example 60% exposure to Dutch and 40% exposure to German.

Input quality: With input quality, I refer to native versus non-native phonetic characteristics in the speech input. For example, monolingual children, who are not exposed to strong dialectal variation, receive their language input from caretakers who are native speakers of the language the child is acquiring. By contrast, the input received by bilingual children is likely to diverge from monolingual children's input because they are often exposed to non-native or attrited speech. In this dissertation, I use the term *input quality* in reference to native versus non-native or attrited productions of VOT by the children's caregivers.

2 Bilingual children's speech production

Perhaps the most intuitive approach to study bilingual children's language acquisition is through their speech production. Technical advancements in the twentieth century enabled objective measures of the speech signal. In particular, phonetic analyses of speech production data became considerably less laborious once personal computers and speech analysis software, such as Praat (developed since 1992; Boersma & van Heuven, 2001; Boersma & Weenink, 2017) became accessible to language researchers. Speech production methods used to study bilingual first language acquisition range from naturalistic speech observations as done by the diarists Ronjat and Leopold to controlled speech elicitations, in which the child produces pre-determined words or utterances (see Kehoe, 2015, for an overview).

A bilingual child's speech can reveal effects of bilingualism, such as the use of both languages in one utterance (code-mixing, e.g., Lanza, 1997; Vihman, 1998), or more narrowly, subtle phonetic influences from one language on the other.

Regarding cross-linguistic influences in the domains of phonetics and phonology, considerable attention has been payed to bilingual children's realization of VOT (Deuchar & Clark, 1996; Fabiano-Smith & Bunta, 2012; Johnson & Wilson, 2002; Kehoe et al., 2004; Khattab, 2000), omission of coda consonants (Almeida, Rose & Freitas, 2012; Ezeizabarrena & Alegria, 2015; Gildersleeve-Neumann, Kester, Davis & Peña, 2008; Keffala, Barlow & Rose, 2018; Lleó, Kuchenbrandt, Kehoe & Trujillo, 2003), and vowel length and quality (Brasileiro, 2009; Kehoe, 2002). These studies yielded diverse results, ranging from monolingual-like productions to complete transfer from one language to the other, supporting that there is considerable variation in bilingual first language acquisition.

In **Chapter 2** of this dissertation, the focus is on bilingual children's production of VOT in Dutch and German in comparison to the VOT production of their monolingual peers, taking into account the bilinguals' input quantity (see section 4 of this chapter below). The results reveal differential acquisition outcomes predominantly in the heritage language German. However, the data in Chapter 2 did not allow drawing firm conclusions regarding the question whether bilingual children have language-specific phonological systems: they distinguished Dutch and German 'voiceless' plosives, but not 'voiced' plosives in their productions. This result led to investigations of the children's voicing *perception* (see section 3 of this chapter below).

In **Chapter 4** of this dissertation, the focus is again on the bilingual children's VOT production in Dutch and German in comparison to their monolingual peers. Dutch contrasts plosives by voicing at the labial (/b/ vs. /p/) and coronal (/d/ vs. /t/) place of articulation, but only has a 'voiceless' dorsal plosive /k/, and is lacking the 'voiced' counterpart phoneme /g/. Through this phonological gap, I explored whether Dutch-speaking children generalize the voicing contrast to the dorsal place of articulation. Moreover, the lack of /g/ in Dutch allowed for testing whether bilingual children's productions of /g/ in German show evidence for cross-linguistic influence from Dutch despite the lack of /g/ in Dutch. The bilingual children's production of /g/ in German appears to be influenced by Dutch, suggesting that features play an important role in children's speech.

3 Bilingual children's speech perception

The seminal work of Eimas, Siqueland, Jusczyk and Vigorito (1971) introduced child language researchers to the discipline of speech perception. Since then, speech perception methods have been widely used to study the early, preverbal stages of language acquisition, which correspond to a developmental time-window that remained inaccessible to researchers before (see Gervain & Mehler, 2010, for a review of speech perception research during the first year of life).

More recently, speech perception studies have been used to study the effect of bilingualism in infants aged between 4 and 15 months (e.g., Albareda-Castellot, Pons & Sebastián-Gallés, 2011; Bosch & Sebastián-Gallés, 2003; Ferjan Ramírez, Ramírez, Clarke, Taulu & Kuhl, 2016; Garcia-Sierra, Rivera-Gaxiola, Percaccio, Conboy, Romo, Klarman, Ortiz & Kuhl, 2011; Liu & Kager, 2016). These studies suggested that infants with bilingual language input are sensitive to phonological contrasts of both languages.

Also in older bilingual children, speech sound perception studies can provide insight into bilingual language development. It is well-known that young children develop native-like perception considerably before they reach the adult-target in production (e.g., Menyuk & Anderson, 1969). It is therefore likely that children deviate from the adult-target in production, but nevertheless show target-like perception.

Speech perception measures are therefore a useful tool to investigate bilingual children's phonological development as they are not restricted by developmental motor constraints that a child may be facing. Yet, the subfield of bilingual children's speech perception is still dominated by infant studies (but see Brasileiro, 2009, and McCarthy, Mahon, Rosen & Evans, 2014, for bilingual preschoolers' speech sound perception). In **Chapter 3** of this dissertation, I use a speech sound perception task to investigate whether bilingual children have language-specific voicing systems, and to determine whether the quantity of their language exposure is associated with their voicing perception. Based on the results, I argue for language-specific voicing systems in bilingual preschoolers, which seem to emerge even when exposure to one language is scarce.

4 Bilingual children's input

Children acquire language through exposure to speech, and a large amount of speech input is usually provided by the parents (Hart & Risley, 1995). While the input *quantity* can be fairly variable even in monolingual language acquisition (De Houwer, 2014; Hart & Risley, 1995; Hoff, 2006), the input of bilingual children may in addition differ from monolingual children's input in *qualitative* aspects (Kehoe, 2015; McCarthy et al., 2014).

The amount of language exposure plays a crucial role in bilingual children's language development (Unsworth, 2013): it is not surprising that a child with very little exposure to one language does not perform on par with a child whose language exposure is more balanced. In order to statistically test the hypothesis that the amount of language exposure is associated with a bilingual child's language acquisition, detailed language background measures and a sample of participants that is large enough to allow for analyses of association are required. These prerequisites pose a challenge to bilingualism research as a large sample of participants with a similar language background is often hard to find, such that the recruitment requires a disproportionate time commitment. A three-year period of data collection enabled me to obtain a reasonably large participant sample for this dissertation to address the effect of input quantity on bilingual first language acquisition. In **Chapter 2**, I explore the effect of the amount of language exposure on bilingual children's VOT production, providing first evidence that more exposure to the heritage language is beneficial for bilingual children's phonological development. The influence of the amount of language exposure on bilingual children's voicing perception is explored in **Chapter 3**, but no association between the two was detected. This suggests that even bilingual children with relatively little heritage language exposure developed language-specific voicing systems in perception.

A qualitative input factor in bilingual language acquisition is the partially non-native language input. When a child is brought up by parents who speak different native languages, or parents who share the native language, but moved to a different country, the child is most likely exposed to at least one parent speaking a non-native language. A vast body of research shows that non-native-like speech productions in (late) second language acquisition are the norm rather than the exception (Flege, 1987, 1991; Flege & Eefting, 1987a, 1987b; Flege & Port, 1981; Simon & Leuschner, 2010). It is therefore highly likely that bilingual children are exposed to foreign-accented speech through their parents in one or possibly both of their languages. A relationship between non-native speech input and bilingual children's divergences from their monolingual peers has previously been suggested, but due to the aforementioned difficulties in recruiting a large and somewhat homogeneous sample of bilingual children, this assumption remained untested to this day (Deuchar & Clark, 1996; Fish, García-Sierra, Ramírez-Esparza & Kuhl, 2017; Khattab, 2003; Klinger, 1962; Mayr & Montanari, 2015). **Chapter 5** of this dissertation focuses on the speech of the bilingual children's parents, and shows that non-native productions of VOT are common in the L2, and even extend to the L1 of the German parents. **Chapter 6** then connects the VOT productions of the mothers to those of their children, providing evidence for a direct relationship between differential maternal VOT and differential child VOT in both Dutch and German.

5 Research questions and dissertation outline

The present dissertation takes a global approach to bilingual preschoolers' phonological development, taking into account their speech production, speech perception, and the quantity and quality of the language input provided by their parents. The five experimental chapters and the main research questions they sought to answer are summarized below.

Chapter 2 forms the base of this dissertation and reports the children's production of VOT in Dutch and German. I establish whether bilingual children's production of VOT is language-specific, and whether bilingual children produce VOT differently than their monolingual peers. I furthermore test whether the age and the current percentage of heritage language exposure of the bilingual children are associated with their VOT production. This chapter provides evidence for partially language-specific productions and differential productions predominantly in the heritage language German. I argue that not age, but the proportion of language exposure is a crucial factor determining bilingual children's VOT production.

Chapter 3 focuses on the children's perception of voicing, and asks whether bilingual children perceive voicing differently in Dutch and in German, and whether the quantity of heritage language input is associated with their voicing perception. In addition, I tested whether bilingual children differ from their monolingual peers in their perception of voicing. The results count in favor of language-specific voicing systems that appear to be independent of proportional language exposure.

Chapter 4 presents new data of the children's production of the 'voiced' dorsal plosive /g/, which exists in German, but not in Dutch. The aim of this chapter is to address whether phonological features play a role in monolingual and bilingual children's speech. I argue that phonological features are crucial to understand the production patterns in children, based on the findings that monolingual and bilingual children generalize the voicing contrast to the dorsal place of articulation, and that bilingual children's productions of /g/ in German appear to be influenced by Dutch.

Chapter 5 examines the speech of the bilingual children's parents in their L1 and L2. The main questions are whether the parents attained native-like VOT production in their L2, and whether they maintained monolingual-like VOT production in their L1. The results provide evidence for differential VOT productions of the German parents in both their L2 Dutch *and* their L1 German.

Chapter 6 connects the findings of chapters 2 and 5 and tests whether the differential VOT productions of the bilingual children's mothers are associated with their children's VOT productions. These results are the first to show that differential language input has measurable effects on bilingual children's speech productions.

Chapter 7 summarizes and discusses the main findings of this dissertation. An overview of possible avenues for future research to further advance the field of bilingualism closes this dissertation.

References

Albareda-Castellot, B., Pons, F., & Sebastián-Gallés, N. (2011). The acquisition of phonetic categories in bilingual infants: New data from an anticipatory eye movement paradigm. *Developmental Science, 14*, 395–401.

Almeida, L., Rose, Y., & Freitas, J. (2012). Prosodic influence in bilingual phonological development: Evidence from a Portuguese-French first language learner. In A. Biller, E. Chung, & A. Kimball (Eds.), *Proceedings of the 36th Annual Boston University Conference on Language Development* (pp. 42–52). Somerville, MA: Cascadilla Press.

Arnberg, L. (1981). *A longitudinal study of language development in four young children exposed to English and Swedish in the home*. Linköping: Linköping University, Department of Education.

Au, T. K., Knightly, L. M., Jun, S.-A., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science, 13*, 238–243.

Boersma, P., & van Heuven, V. (2001). Speak and unSpeak with Praat. *Glot International, 5*, 341–347.

Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer. Version 6.0.28. Retrieved from http://www.praat.org/

Bosch, L., & Sebastián-Gallés, N. (1997). Native language recognition abilities in four-month-old infants from monolingual and bilingual environments. *Cognition, 65*, 33–69.

Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech, 46*, 217–243.

Brasileiro, I. (2009). *The effects of bilingualism on children's perception of speech sounds.* Utrecht: LOT.

Cummins, J. (2005). A proposal for action: Strategies for recognizing heritage language competence as a learning resource within the mainstream classroom. *Modern Language Journal, 89,* 585–592.

De Houwer, A. (1990). *The acquisition of two languages from birth: A case study.* Cambridge: University Press.

De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics, 28,* 411–424.

De Houwer, A. (2009). *Bilingual first language acquisition*. Bristol: Multilingual Matters.

De Houwer, A. (2011). Language input environments and language development in bilingual acquisition. *Applied Linguistics Review, 2*, 221–240.

De Houwer, A. (2014). The absolute frequency of maternal input to bilingual and monolingual children: A first comparison. In T. Grüter & J. Paradis (Eds.), *Input and experience in bilingual development* (pp. 37–58). Amsterdam: Benjamins.

De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2014). A bilingual-monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, *35*, 1189–1211.

Deuchar, M., & Clark, A. (1996). Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics, 24,* 351–365.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science,* 171, 303–306.

Ezeizabarrena, M.-J., & Alegria, A. (2015). Early coda production in bilingual Spanish and Basque. In T. Judy & S. Perpiñán (Eds.), *The acquisition of Spanish in understudied language pairings* (pp. 75–104). Amsterdam: Benjamins.

Fabiano, L., & Goldstein, B. (2005). Phonological transfer in bilingual Spanish-English speaking children. *Journal of Multilingual Communication Disorders, 3,* 56–63.

Fabiano-Smith, L., & Bunta, F. (2012). Voice onset time of voiceless bilabial and velar stops in 3-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics & Phonetics, 26,* 148–163.

Ferjan Ramírez, N., Ramírez, R. R., Clarke, M., Taulu, S., & Kuhl, P. K. (2016). Speech discrimination in 11-month-old bilingual and monolingual infants: A magneto-encephalography study. *Developmental Science, 20*, 1–16.

Fikkert, P. (2010). Developing representations and the emergence of phonology: Evidence from perception and production. In C. Fougeron, B. Kuehnert, M. Imperio & N. Vallee (Eds.), *Laboratory Phonology 10* (pp. 227–260). Berlin: De Gruyter Mouton.

Fish, M. S., García-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. (2017). Infant-directed speech in English and Spanish: Assessments of monolingual and bilingual caregiver VOT. *Journal of Phonetics, 63*, 19–34.

Flege, J. E. (1987). The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*, 47–64.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, *89*, 395–411.

Flege, J. E., & Eefting, W. (1987a). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, *15*, 67–83.

Flege, J. E., & Eefting, W. (1987b). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication, 6*, 185–202.

Flege, J. E., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech, 24*, 125–146.

Garcia-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., Ortiz, S., & Kuhl, P. K. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. *Journal of Phonetics, 39*, 546–557.

Gathercole, V., Thomas, E., Roberts, E., Hughes, C., & Hughes, E. (2013). Why assessment needs to take exposure into account: Vocabulary and grammatical abilities in bilingual children. In V. Gathercole (Ed.), *Issues in the assessment of bilinguals* (pp. 20–55). Bristol: Multilingual Matters.

Genesee, F. (1989). Early bilingual development: One language or two? *Journal of Child Language*, *16*, 161–179.

Genesee, F., & Nicoladis, E. (2005). Bilingual acquisition. In E. Hoff & M. Shatz (Eds.), *Handbook of Language Development* (pp. 324–342). Oxford: Blackwell.

Gervain, J., & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual Review of Psychology*, *61*, 191–218.

Gildersleeve-Neumann, C., Kester, E., Davis, B., & Peña, E. (2008). English speech sound development in preschool-aged children from bilingual English-Spanish environments. *Language, Speech, and Hearing Services in Schools, 39*, 314–328.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*, 3–15.

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Brookes.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*, 55–88.

Johnson, C., & Wilson, I. (2002). Phonetic evidence for early language differentiation: Research issues and some preliminary data. *International Journal of Bilingualism, 6*, 271–289.

Keffala, B., Barlow, J. A., & Rose, S. (2018). Interaction in Spanish-English bilinguals' acquisition of syllable structure. *International Journal of Bilingualism, 22,* 16–37.

Kehoe, M. (2002). Developing vowel systems as a window to bilingual phonology. *International Journal of Bilingualism, 6,* 315–334.

Kehoe, M. (2015). Cross-linguistic interaction: A retrospective and prospective view. In E. Babatsouli & D. Ingram (Eds.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2015* (pp. 141–167). Retrieved from http://ismbs.eu/data/documents/Proceedings-ISMBS-2015.pdf

Kehoe, M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition, 7,* 71–88.

Kellerman, E., & Sharwood Smith, M. (1986). *Cross-linguistic influence in second language acquisition*. Elmsford, NY: Pergamon.

Khattab, G. (2000). VOT in English and Arabic bilingual and monolingual children. In D. Nelson & P. Foulkes (Eds.), *Leeds Working Papers in Linguistics and Phonetics, 8* (pp. 95–122). Retrieved from http://www.leeds.ac.uk/arts/homepage/387/leeds_working_papers_in_linguistics_and_phonetics

Khattab, G. (2003). Age, input, and language mode factors in the acquisition of VOT by English-Arabic bilingual children. In M. J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the International Congress of Phonetic Sciences, 15* (pp. 3213–3216). Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_3213.pdf

Klinger, H. (1962). Imitated English cleft palate speech in a normal Spanish speaking child. *Journal of Speech and Hearing Disorders, 27,* 379–381.

Kupisch, T. (2013). A new term for a better distinction? A view from the higher end of the proficiency scale. *Theoretical Linguistics, 39*, 203–214.

Kupisch, T., Akpinar, D., & Stöhr, A. (2013). Gender assignment and gender agreement in adult bilinguals and second language learners of French. *Linguistic Approaches to Bilingualism, 3,* 150–179.

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 the weaker and stronger language. *Journal of French Language Studies, 24*, 347–376.

Kupisch, T., & Rothman, J. (2016). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*. Advance online publication. doi: 10.1177/1367006916654355

Lambert, W. E. (1981). Bilingualism and language acquisition. *Annals of the New York Academy of Sciences, 379*, 9–22.

Lanza, E. (1997). *Language mixing in infant bilingualism: A sociolinguistic perspective*. Oxford: Clarendon.

Leopold, W. F. (1939-1949). *Speech development of a bilingual child*. Evanston, IL: Northwestern University Press.

Levelt, W. J. M. (2014). *A history of psycholinguistics. The pre-Chomskyan era.* Oxford: University Press.

Liu, L., & Kager, R. (2016). Perception of a native vowel contrast by Dutch monolingual and bilingual infants: A bilingual perceptual lead. *International Journal of Bilingualism, 20*, 335–345.

Lleó, C., & Kehoe, M. (2002). On the interaction of phonological systems in child bilingual acquisition. *International Journal of Bilingualism, 6,* 233–237.

Lleó, C., Kuchenbrandt, I., Kehoe, M., & Trujillo, C. (2003). Syllable final consonants in Spanish and German monolingual and bilingual acquisition. In N. Müller (Ed.), *(Non)Vulnerable Domains in Bilingualism* (pp. 191–220). Amsterdam: Benjamins.

Mayr, R., & Montanari, S. (2015). Cross-linguistic interaction in trilingual phonological development: The role of the input in the acquisition of the voicing contrast. *Journal of Child Language, 42,* 1006–1035.

McCarthy, K. M., Mahon, M., Rosen, S., & Evans, B. G. (2014). Speech perception and production by sequential bilingual children: A longitudinal study of voice onset time acquisition. *Child Development, 85*, 1965–1980.

Meisel, J. (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.

Menyuk, P., & Anderson, S. (1969). Children's identification and reproduction of /w/, /r/, and /l/. *Journal of Speech, Language, and Hearing Research, 12,* 39–52.

Montrul, S. (2016). *The acquisition of heritage languages.* Cambridge: University Press.

Müller, N., & Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition, 4*, 1–21.

Nicoladis, E. & Genesee, F. (1997). Language development in preschool bilingual children, *Journal of Speech-Language Pathology and Audiology, 21*, 258–270.

Oller, D. K., Eilers, R. E., Urbano, R., & Cobo-Lewis, A. B. (1997). Development of precursors to speech in infants exposed to two languages. *Journal of Child Language, 24*, 407–425.

Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition, 18,* 1–25.

Pearson, B. Z., Fernández, S. C., Lewedag, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants (ages 10 to 30 months). *Applied Psycholinguistics, 18,* 41–58.

Ronjat, J. (1913). *Le développement du langage observé chez un enfant bilingue.* Paris: Champion.

Rothman, J., & Treffers-Daller, J. (2014). A prolegomenon to the construct of the native speaker: Heritage speaker bilinguals are natives too! *Applied Linguistics, 35*, 93–98.

Simon, E., & Leuschner, T. (2010). Laryngeal systems in Dutch, English, and German: A contrastive phonological study on second and third language acquisition. *Journal of Germanic Linguistics, 22*, 403–424.

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*, 86–110.

Valdés, G. (2000a). Introduction. In Spanish for native speakers. *AATSP professional development series handbook for teachers K-16, Volume 1* (pp. 1–20). New York, NY: Harcourt College.

Valdés, G. (2000b). Teaching heritage languages: An introduction for Slavic language-teaching professionals. In O. Kagan & B. Rifkin (Eds.), *Learning and teaching of Slavic languages and cultures: Toward the 21st century* (pp. 375–403). Bloomington, IN: Slavica.

Vihman, M. (1998). A developmental perspective on codeswitching: Conversations between a pair of bilingual siblings. *International Journal of Bilingualism, 2,* 45–84.

Volterra, V., & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language, 5*, 311–326.

Werker, J. F., & Byers-Heinlein, K. (2008). Bilingualism in infancy: First steps in perception and comprehension. *Trends in Cognitive Science, 12*, 144–151.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development, 1*, 197–234.

Chapter 2

Heritage language exposure impacts voice onset time of Dutch-German simultaneous bilingual preschoolers

2

Based on:

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017). Heritage language exposure impacts voice onset time of Dutch-German simultaneous bilingual preschoolers. *Bilingualism: Language and Cognition*. Advance online publication. doi: 10.1017/S1366728917000116

Abstract

This study assesses the effects of age and language exposure on VOT production in 29 simultaneous bilingual children aged 3;7 to 5;11 who speak German as a heritage language in the Netherlands. Dutch and German have a binary voicing contrast, but the contrast is implemented with different VOT values in the two languages. The results suggest that bilingual children produce 'voiced' plosives similarly in their two languages, and these productions are not monolingual-like in either language. Bidirectional cross-linguistic influence between Dutch and German can explain these results. Yet, the bilinguals seemingly have two autonomous categories for Dutch and German 'voiceless' plosives. In German, the bilinguals' aspiration is not monolingual-like, but bilinguals with more heritage language exposure produce more target-like aspiration. Importantly, the amount of exposure to German has no effect on the majority language's 'voiceless' category. This implies that more heritage language exposure is associated with more language-specific voicing systems.

1 Introduction

Bilingual children's realization of the voicing contrast has received substantial attention in language acquisition research during the past two decades, and consistently revealed differences from monolingual children's VOT production (Deuchar & Clark, 1996; Fabiano-Smith & Bunta, 2012; Johnson & Wilson, 2002; Kehoe, Lleó & Rakow, 2004; Khattab, 2000; McCarthy, Mahon, Rosen & Evans, 2014). These studies have been conducted mainly on small samples of bilinguals immersed in an aspiration language (i.e., English, except for Kehoe et al., 2004 on German) with a prevoicing language as the minority language. Although these studies used adequate statistical analyses, they were not designed to statistically assess the effects of age or language exposure, which are important factors in monolingual and bilingual language acquisition (Armon-Lotem & Ohana, 2017; Gathercole & Hoff, 2007; Gathercole & Thomas, 2009; Mayr & Siddika, 2016; Unsworth, 2013; Yu, De Nil & Pang, 2015). To determine to what extent age and language exposure can explain bilinguals' linguistic behaviors, samples of participants must be large enough to allow for association analyses. Furthermore, it is essential to the field of early bilingual phonological acquisition to determine whether previous findings on minority languages acquired in an English-dominant environment extend to other acquisition settings and languages (Kehoe, 2015). The present study is the first to address these outstanding issues in a sample of Dutch-German bilingual preschoolers that is large enough to allow for association analyses between the effects of both age and language exposure, and bilingual children's speech production.

Simultaneous bilingual children acquire two native languages from birth or shortly thereafter. From then on, the two languages are accommodated in their brain and are likely to influence each other, a phenomenon known as CROSS-LINGUISTIC INFLUENCE (CLI; Fabiano & Goldstein, 2005; Fabiano-Smith & Bunta, 2012; Kehoe, 2002; Kehoe et al., 2004; Kellerman & Sharwood Smith, 1986; Lleó & Kehoe, 2002; Müller & Hulk, 2001; Paradis & Genesee, 1996). It has been well documented that bilinguals who acquire their second language (L2) at a later age (denoted as sequential bilinguals) are often affected by CLI from first language (L1) to L2 phonology (e.g., Flege, 1991; Flege & Port, 1981; Laeufer, 1996; Williams, 1980). Much less is known about the impact of CLI on phonological development in young simultaneous bilingual children (Deuchar & Clark, 1996; Fabiano-Smith & Bunta, 2012; Johnson & Wilson, 2002; Kehoe et al., 2004).

CLI can cause bilingual speech to be differential or not 'native-like' (see Kupisch & Rothman, 2016 for a critical perspective on terminology), meaning that bilinguals produce speech sounds differently from monolinguals. When a bilingual's speech differs from a monolingual's speech, it may be perceived as foreign-accented (Flege, 1984; Major, 1987; Riney & Takagi, 1999; Sancier & Fowler, 1997; Schoonmaker-Gates, 2015). Such differential bilingual speech can still be 'language-specific' if similar sounds are produced differently in the two languages.

Conversely, CLI may have facilitative effects on bilinguals' language development and accelerate their acquisition of certain linguistic structures compared to monolingual acquisition (Grech & Dodd, 2008; Mayr, Howells & Lewis, 2015; Tamburelli, Sanoudaki, Jones & Sowinska, 2015). Acceleration can occur when one of the bilingual's languages contains a difficult and/or infrequent structure that is more frequent in the other language. The practice with such a structure in one language may have facilitative effects in the other language.

Bilinguals acquire two languages in the same amount of time in which a monolingual acquires a single language, resulting in overall less exposure and therefore less experience with each language relative to monolingual acquisition (Gathercole & Thomas, 2005, 2009; Unsworth, 2008; Unsworth, Argyri, Cornips, Hulk, Sorace & Tsimpli, 2014). Reduced exposure likely results in slower acquisition of linguistic structures that are distinct between the bilingual child's two languages. As a result of this reduced exposure, bilinguals may reach certain developmental stages later than their age-matched monolingual peers.

To date, there is no framework that specifically targets the speech of young simultaneous bilingual children. However, models of the speech of sequential bilingual adults and monolingual children are available and can be extended to account for CLI and language exposure effects in simultaneous bilingual children. The SPEECH LEARNING MODEL (SLM; Flege, 1995) originally focuses on age of acquisition-related constraints on native-like production of L2 sounds, and can partially account for CLI in simultaneous bilinguals' speech (Fabiano-Smith & Bunta, 2012; Fabiano-Smith & Goldstein, 2010; Gildersleeve-Neumann & Wright, 2010). The SLM assumes that many production errors in the L2 are rooted in sound perception, and puts forward seven hypotheses of the L2 learner's sound perception, sound processing and storage, and sound production. Two of these hypotheses can be extended to the sound production of simultaneous bilingual children.

The first hypothesis, henceforth the 'Age of Acquisition Hypothesis', states that increasing age of acquisition goes hand in hand with a decreasing ability to distinguish L1 and L2 sounds. This hypothesis inversely suggests that an early age of acquisition promotes the ability to discriminate between sounds, resulting in less CLI and more language-specific acquisition of speech sounds. In the case of simultaneous bilingual acquisition, both languages are acquired in parallel from birth, and the Age of Acquisition Hypothesis can be extended to suggest that simultaneous bilingual children may be less prone to CLI and are likely to acquire monolingual-like sounds in both of their languages.

The second hypothesis, henceforth the 'Equivalence Classification Hypothesis' (see also Flege, 1987) formulates an exception to the Age of Acquisition Hypothesis. Equivalence classification is one form of CLI and proposes that the formation of new phonological categories may be blocked if an L2 sound overlaps with a similar L1 position-sensitive allophone. In the context of simultaneous bilingual acquisition, equivalence classification may cause a bilingual child to acquire only one category for two sounds that she perceives to be alike in the two languages. Such category mergers are natural language change processes that normally unfold over time in language communities (Romaine, 1978; Wells, 1982). In sum, the SLM can account for differential sound production by simultaneous bilinguals as a result of CLI in the perception and category formation of sounds that are perceptually similar between the two languages. This model does not ascribe the bilinguals' differential sound production to differences in language exposure between bilinguals and monolinguals.

The second model that can be extended to the speech of young bilinguals is the A(RTICULATORY)-MAP model (McAllister Byun, Inkelas & Rose, 2016), which explains differences between (monolingual) child and adult speech through anatomical and motor control differences. The model proposes that experience-based information about previous articulator movements and the resulting acoustic output is stored in episodic memory. Two grammatical constraints draw on these episodic traces: ACCURACY formalizes the pressure to match adult speech production, while PRECISION formalizes the pressure to produce stable and well-practiced realizations, even if they do not perfectly match the adult-target. Interactions between accuracy, precision, and other relevant constraints, such as developmental constraints, determine a child's actual speech production. The A-Map model explicitly predicts that children's speech production becomes increasingly precise with more production experience, leading to a decreasing deviation from the adult-target. Bilingual children necessarily gain less production experience than monolinguals with sounds that occur in only one of their languages.

The A-Map model extended to bilingual children can account for delays in bilinguals' production of articulatory complex sounds that are limited to one of their languages. The bilinguals' reduced production experience in combination with the precision constraint explains that bilinguals take longer than age-matched monolinguals to reach the adult-target for such sounds. However, bilingual children may gain more production experience than monolinguals with sounds that exist in both the bilingual's languages, but with differing frequency. In these cases, the bilingual A-Map encompasses more traces of motor-actions and acoustic outcomes in episodic memory than the monolingual A-Map, which may accelerate target-like production of that structure. In sum, the A-Map model extended to simultaneous bilingual children's speech offers a framework that captures how different production experience across two languages delays the acquisition of unshared speech sounds. Linked to production experience, the extended A-Map model can also account for acceleration effects in bilinguals' speech through motor practice accumulated in the other language, which can be interpreted as positive CLI.

Irrespective of these theoretical models, disentangling CLI and language exposure as possible reasons for linguistic differences between bilinguals and monolinguals is inherently difficult because acquiring two languages necessarily reduces the exposure to each language. It is possible, however, to assess language exposure effects by relating linguistic differences within a bilingual population to individual differences in language exposure – provided the sample is large enough to allow for association analyses. Once the exposure effects have been assessed, one can establish which findings require an additional explanation in terms of CLI. The present study addressed these issues with regards to VOICE ONSET TIME (VOT).

1.1 Voice onset time

Voice onset time (VOT) is an acoustic cue that contributes to the phonological distinction between 'voiced' and 'voiceless' plosives, such as /b/ and /p/. VOT is the duration of the interval between the start of vocal fold vibration relative to the release of a plosive's burst, and is the most important cue to voicing (Abramson & Lisker, 1973; Cho & Ladefoged, 1999; van Alphen, 2004; van Alphen & Smits, 2004). Although many of the world's languages have a two-way contrast[1] between 'voiced' and 'voiceless' plosives, this phonological contrast can have different phonetic implementations. As schematized in Figure 1, the VOT continuum can be divided into three phonetic categories: prevoicing, short lag, and aspiration. Languages like Dutch, Arabic, French, Japanese, Spanish, and Sylheti contrast 'voiced' and 'voiceless' plosives by means of prevoicing vs. short lag VOT. Languages like

[1] Eastern Armenian and Thai have a three-way voicing contrast (Lisker & Abramson, 1964), and East Bengali, Hindi and Nepali have a four-way voicing contrast (Davis, 1995; Mikuteit & Reetz, 2007; Poon & Mateer, 1985).

German and English implement the voicing contrast with short lag VOT vs. aspiration[2]. Language-specific VOT values within these ranges may differ cross-linguistically.

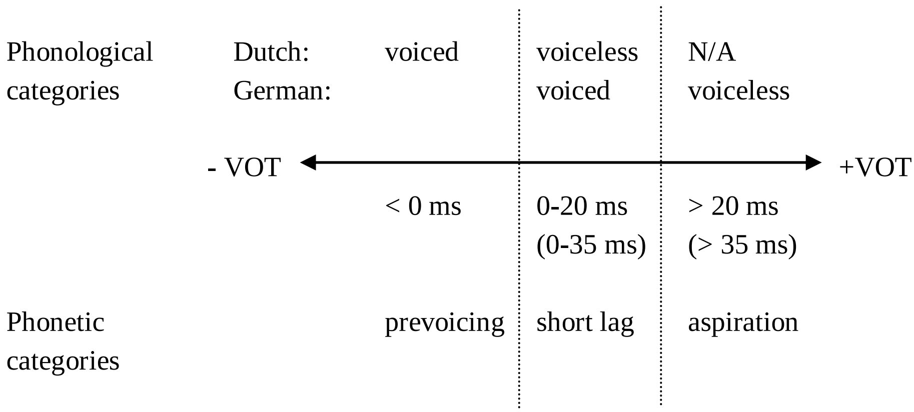


Figure 1. The VOT continuum: phonological and phonetic categories.

The 0 ms point in a VOT continuum denotes the plosive's burst release. Vocal fold vibration that starts prior to burst release falls into the prevoicing range. Prevoiced plosives are phonologically and phonetically described as 'voiced', and occur for example in Dutch (Deighton-van Witsen, 1976; Lisker & Abramson, 1964; van Alphen & Smits, 2004). If the onset of voicing falls between 0 ms and approximately 20–35 ms after the burst release, the plosive falls within the short lag VOT range. Phonetically, such sounds can be described as devoiced, but phonologically, they can be classified as 'voiceless' or 'voiced', depending on the language. In Dutch, plosives produced with short lag VOT are considered the 'voiceless' counterpart of prevoiced plosives. In other languages, like German, short lag plosives represent the majority of 'voiced' plosives. Although not required in German, adults sometimes prevoice even up to around 50% of their 'voiced' plosives (Fischer-Jørgensen, 1976; Hamann & Seinhorst, 2016; Jessen, 1998; Kohler, 1977; Stock, 1971).

If the onset of voicing exceeds the 20–35 ms upper limit of short lag VOT, the plosive falls within the aspiration range on the VOT continuum. These aspirated plosives are always phonologically 'voiceless' and represent the 'voiceless' counterparts to 'voiced' short lag plosives in German. The duration of aspiration

[2] Swedish distinguishes prevoiced 'voiced' and aspirated 'voiceless' plosives (Beckman, Helgason, McMurray & Ringen, 2011).

typically averages between 45–70 ms in adult native speakers of German (Fischer-Jørgensen, 1976; Haag, 1979; Jessen, 1998; Neuhauser, 2011).

Even though we construe the three VOT ranges – prevoicing, short lag, and aspiration – as relatively fixed, small VOT differences within each range can arise due to language-internal factors. VOT generally increases the further the place of articulation is to the back of the mouth (Fischer-Jørgensen, 1954; Lisker & Abramson, 1964; Maddieson, 1997; Nearey & Rochet, 1994; Peterson & Lehiste, 1960; Umeda, 1977; van Alphen & Smits, 2004; Volaitis & Miller, 1992). In addition, word-initial aspirated plosives have longer VOT when they occur in monosyllabic as opposed to polysyllabic words, but VOT in short lag and prevoiced plosives seems to be unaffected by word length (Flege, Frieda, Walley & Randazza, 1998; Yu et al., 2015). Short lag and aspirated plosives that appear before close vowels tend to be produced with longer VOT than plosives followed by open vowels (Nearey & Rochet, 1994; Yeni-Komshian, Caramazza & Preston, 1977). Speaking rate further influences VOT in continuous speech: at a fast speaking rate, the duration of aspiration and prevoicing decreases (Kessinger & Blumstein, 1997).

1.2 VOT development in monolingual children

Monolingual children start to produce short lag plosives in their early babbles, irrespective of whether their native language contrasts voicing by means of short lag VOT and aspiration or prevoicing and short lag VOT (Eilers, Oller & Benito-Garcia, 1984; Kager, van der Feest, Fikkert, Kerkhoff & Zamuner, 2007; Kewley-Port & Preston, 1974; Macken & Barton, 1980a; Oller & Eilers, 1982; Oller, Wieman, Doyle & Ross, 1976; Zlatin & Koenigsknecht, 1976). Research on aspiration development revealed that children reliably produce aspiration around the second birthday (Eilers et al., 1984; Kager et al., 2007; Kewley-Port & Preston, 1974; Macken & Barton, 1980a; Oller & Eilers, 1982; Oller et al., 1976; Zlatin & Koenigsknecht, 1976). Children start to produce adult-like prevoicing later in life, possibly during the early school years (Allen, 1985; Bortolini, Zmarich, Fior & Bonifacio, 1995; Kager et al., 2007; Khattab, 2000; Macken & Barton, 1980b; MacLeod, 2017).

Research on the acquisition of aspiration found that English-speaking children between 0;6 and 4;6 develop a voicing contrast by 2;6, which is similar to the contrast of older children, but not yet adult-like (Kewley-Port & Preston, 1974). Longitudinal data from English-speaking children starting at age 1;6 to just after 2;0 revealed three acquisition stages (Macken & Barton, 1980a): 1) ‘voiced’ and 'voiceless' plosives have short lag VOT; 2) 'voiced' and 'voiceless' plosives have a covert contrast within the short lag range that is presumably not perceived by adults; and 3) 'voiceless' plosives have adult-like aspiration. Other research found that English-speaking two-year-olds (2;6–3;0) and six-year-olds (6;1–6;11) produce on average shorter aspiration in 'voiceless' plosives than adults despite producing an overt and reliable voicing contrast (Zlatin & Koenigsknecht, 1976). Data on languages other than English are sparse, but one case study showed that a German-speaking child aged 1;0 to 2;2 initially aspirated 50% of 'voiceless' plosives and only reliably aspirated by age 2;0 (Kager et al., 2007). The finding that children commonly produce aspiration values diverging from adults can be related to still-developing control of timing between the plosive's burst release and the onset of vocal fold vibration (Barton & Macken, 1980; Kewley-Port & Preston, 1974; Koenig, 2000; Macken & Barton, 1980a; Menyuk & Klatt, 1975; Whiteside, Dobbin & Henry, 2003; Yu et al., 2015; Zlatin & Koenigsknecht, 1976). In sum, children acquiring an aspiration language overtly distinguish 'voiceless' from 'voiced' plosives by approximately two years of age, although the length of aspiration may still be different from adults.

Research on the acquisition of prevoicing found that Dutch-speaking children aged between 1;0 and 1;2 prevoice only 30% of all 'voiced' plosives. The percentage of prevoiced 'voiced' plosives increases to 60% by the end of their third year of life (Kager et al., 2007). The majority of Italian-speaking children aged between 1;6 and 1;9 do not contrast plosives by voicing and instead produce the majority of plosives within the short lag VOT range (Bortolini et al., 1995). French-speaking children aged between 1;9 and 2;8 generally avoid 'voiced' plosives and prevoice less than 2% of all produced plosives (Allen, 1985). Longitudinal data of Spanish-speaking children aged 1;7 to 2;1 and at 3;10 revealed that even at the age of almost 4, children still do not reliably produce prevoicing for 'voiced' plosives (Macken & Barton, 1980b). Instead, 'voiced' plosives are spirantized – that is, produced as fricatives – to make a voicing distinction. Between 2;6 and 4;6, Canadian French-speaking children acquire a voicing contrast that nevertheless differs phonetically from adult ranges in that they produce prevoicing less reliably than adults (MacLeod, 2017). Arabic-speaking children produce prevoicing inconsistently at 5;4 and even 7;4, but seem to have acquired adult-like prevoicing at 10;3 (Khattab, 2000). In sum, prevoicing poses a challenge to young children and non-target-like production persists in school-aged children. Table 1 summarizes details about the studies on monolingual children's VOT development.

Table 1. Studies on monolingual children's VOT development.

Voicing contrast	Study	Language	Age	*N*
Short lag / aspiration	Kewley-Port & Preston (1974)	English	0;6–2;5, 3;6, 4;6	10
	Macken & Barton (1980a)	English	1;6–2;0	4
	Zlatin & Koenigsknecht (1976)	English	2;6–3;0, 6;1–6;11	20
	Kager et al. (2007)	German	1;0–2;2	1
Prevoicing / short lag	Kager et al. (2007)	Dutch	1;0–2;11	11
	Bortolini et al. (1995)	Italian	1;6–1;9	14
	Allen (1985)	French	1;9–2;8	6
	Macken & Barton (1980b)	Spanish	1;7–2;1, 3;10	7
	MacLeod (2017)	French	2;6–4;6	63
	Khattab (2000)	Arabic	5;4, 7;4, 10;3	3

1.3 VOT development in bilingual children

Bilingual children who simultaneously acquire a prevoicing language like Dutch and an aspiration language like German have to acquire plosive categories from both languages. They further need to resolve the phonological ambiguity of the short lag VOT range that corresponds to 'voiceless' plosives in Dutch, and to 'voiced' plosives in German.

During the last two decades, researchers turned to the question how children's VOT develops when they grow up with two languages that differ in their implementation of voicing (Deuchar & Clark, 1996; Fabiano-Smith & Bunta, 2012; Johnson & Wilson, 2002; Kehoe et al., 2004; Khattab, 2000; Mayr & Siddika, 2016; McCarthy et al., 2014; Table 2 provides an overview of the investigated languages, environments and participants). All these studies report on the acquisition of a majority language that has aspiration and a heritage language that has prevoicing, and most report data of the bilinguals' two languages. The results are variable, as will be discussed in more detail below, with a general emergent pattern that aspiration is acquired early and that prevoicing is generally avoided, which resembles the monolingual acquisition pattern.

Deuchar and Clark (1996) investigated a bilingual English-Spanish speaking child in England recorded at 1;7, 1;11 and 2;3. During this period, the child acquired the English voicing distinction between short lag VOT and aspiration, but produced only short lag plosives in Spanish, which is similar to monolingual Spanish-learning children of this age. Khattab (2000) reported data from three bilingual English-Arabic speaking children in England aged 5;6, 7;1 and 10;2 and three age-matched monolingual children in each language. Although the children were older than the one in Deuchar and Clark (1996), their VOT pattern was similar.

Table 2. Studies on bilingual children's VOT development.

Study	Prevoicing language	Aspiration language	Majority language	Age	*N*
Deuchar & Clark (1996)	Spanish	English	English	1;7–2;3	1
Khattab (2000)	Arabic	English	English	5;6–10;2	3
Johnson & Wilson (2002)	Japanese	English	English	2;10–4;11	2
Mayr & Siddika (2016)	Sylheti	English	English	3;7–5;0	20
McCarthy et al. (2014)	Sylheti*	English	English	4;4, 5;4	40
Fabiano-Smith & Bunta (2012)	Spanish	English	Engl. + Span.	3;0–3;11	8
Kehoe et al. (2004)	Spanish	German	German	1;0–3;0	4

*Sylheti was spoken by the children but not explicitly examined in this study.

In English, the bilingual children produced VOT values similar to monolinguals. In Arabic, two of the three bilingual children did not produce prevoicing for 'voiced' plosives, but inconsistent prevoicing was also observed in the five- and seven-year-old Arabic-speaking monolinguals. Johnson and Wilson (2002) recorded two bilingual English-Japanese speaking children in Canada at 2;10 and 3;0 for one child and at 4;8 and 4;11 for the other child. Both children produced aspirated 'voiceless' plosives and short lag 'voiced' plosives in English. Unlike the bilinguals of Deuchar and Clark (1996) and Khattab (2000), the bilinguals contrasted voicing in their heritage language Japanese, but with an English-like contrast between short lag VOT and aspiration. The older child produced longer VOT for /p/ and /t/ in English than in Japanese, but no evidence for language differentiation was observed in the younger child. Similar findings come from Mayr and Siddika (2016) who investigated VOT of 20 Sylheti-English speaking bilingual children aged 3;7 to 5;0 in Wales (10 second-generation bilinguals and 10 third-generation bilinguals). In English, both groups of children produced target-like VOT. In Sylheti, both groups produced 'voiceless' plosives with aspiration, and most 'voiced' plosives with short lag VOT. Only the second-generation bilinguals produced some 'voiced' plosives with prevoicing. Yet, the children's Sylheti VOT was not entirely English-like: the second-generation bilinguals produced longer VOT in English /k, g, t/, and the third-generation bilinguals produced longer VOT in English /k/. In a longitudinal study, McCarthy et al. (2014) investigated the acquisition of English VOT in 40 sequential bilingual Sylheti-English speaking children in England and 15 monolingual English-speaking children. At the first time of testing, the bilinguals had been exposed to English for an average of 7 months. Their English VOT in labial and dorsal plosives was tested at about age 4;4 and 5;4. In line with the findings of Deuchar and Clark (1996), Mayr and Siddika (2016), Khattab (2000), and Johnson and Wilson (2002), the bilinguals produced VOT for English ‘voiceless’ plosives similar to monolinguals in both testing sessions. The bilinguals’ VOT for English ‘voiced’ plosives was significantly shorter than that of monolinguals in the first testing session, but became indistinguishable from monolinguals’ VOT in the second testing session. These five studies indicate that the acquisition of aspiration is not problematic in bilingual acquisition when the children are immersed in a country in which the aspiration language is the majority language. CLI from the aspiration of the majority language to the heritage language may occur (Johnson & Wilson, 2002; Mayr & Siddika, 2016). CLI of prevoicing from the minority language can also play a role, at least in the Sylheti-English speaking sequential bilinguals in McCarthy et al. (2014), and this has similarly been shown for older child L2-learners (Heselwood & McChrystal, 2000).

The studies discussed so far originated from English-speaking countries where English was the medium of instruction at daycare and school, while the use of the heritage language was mostly limited to the home-context. Only the children in McCarthy et al. (2014) were regularly exposed to their heritage language in the London-Bengali community. The acquisition process is potentially different in an environment in which exposure to both languages is more balanced, with frequent input from multiple speakers and schooling in both languages. Fabiano-Smith and Bunta (2012) evaluated VOT of /p/ and /k/ in eight Spanish-English speaking bilingual children aged 3;0 to 3;11 in a Spanish-speaking immigrant community in the United States, where they attended a bilingual preschool. Although the children were raised in the United States, their broader environment provided them with frequent language input from multiple speakers in both English and Spanish. The bilinguals’ productions were compared to those of eight age-matched monolinguals per language. Interestingly, the bilinguals’ VOT pattern was different from the studies described above, in which heritage language exposure was mostly limited to the home context. In English, the bilinguals of Fabiano-Smith and Bunta (2012) produced overall shorter – and thus more Spanish-like – VOT than monolinguals, although this difference was only statistically significant for /k/. In Spanish, no VOT differences were observed between bilinguals and monolinguals. In addition, there was no evidence for VOT differentiation between the bilinguals’ two languages. This study suggests that aspiration can be prone to delayed or differential acquisition in bilinguals when the aspiration language does not provide the clear majority of children’s input. In addition, CLI from Spanish to English can explain the shorter, more Spanish-like, VOT in English.

Individual differences between bilingual children can account for different patterns of VOT development even in similar acquisition contexts. Kehoe et al. (2004) investigated VOT production of four bilingual German-Spanish speaking children in Germany and three monolingual German-speaking children. Recordings took place every other week starting when the children began producing words (1;0 to 1;3) through to approximately 2;6 to 3;0 years. The four bilingual children reflected three different patterns of VOT development: delay, transfer (CLI), and autonomously developing systems. Two bilingual children showed a delay in their VOT development, as they had not acquired a target-like voicing contrast in German by the end of data collection. One bilingual child showed evidence for bidirectional CLI with instances of prevoicing in German and aspiration in Spanish. Nevertheless, the child maintained a distinction between German and Spanish VOT (see also Johnson & Wilson, 2002; Mayr & Siddika, 2016). The fourth bilingual child showed no evidence for CLI. By 2;3 to 2;6, he acquired a voicing opposition between short lag VOT and aspiration in German. Similar to monolingual Spanish acquisition, no voicing opposition had been acquired in Spanish, and instead ‘voiced’ and ‘voiceless’ plosives were both produced with short lag VOT (see also Deuchar & Clark, 1996; Khattab 2000).

In sum, previous work on the acquisition of VOT in young bilingual children demonstrated that the phonologies of bilinguals often interact in a way that can be interpreted as CLI. However, Khattab (2000) emphasizes that the absence of prevoicing in the heritage language is not necessarily related to CLI from the majority language, but may be due to insufficient heritage language exposure.

The review above also revealed variability in bilingual children’s patterns of VOT development in seemingly similar acquisition contexts. A possible reason for these different developmental patterns may be rooted in individual variation in the amount of language exposure (Mayr & Siddika, 2016). Due to relatively small sample sizes, previous research did not allow to statistically test the role of individual differences in language exposure on VOT development. Further, all studies had been conducted in countries where the majority language had aspiration, which brings into question whether similar acquisition patterns are observed when the prevoicing language is the majority language. The current study is designed to address these still outstanding issues.

1.4 The current study

The current study investigates VOT production of Dutch-German speaking simultaneous bilingual children aged 3;7 to 5;11 in the Netherlands who acquired German from one or both parents from birth. This study is the first to investigate effects of age and relative language exposure on VOT production of bilingual children. In contrast to previous research in which the majority language was an aspiration language, the children in this study are immersed in a prevoicing language (Dutch). In addition, Dutch and German monolingual children were tested in the same experimental paradigm. First, we verify the expected VOT production differences between monolingual Dutch and German preschoolers. We then turn to the following three research questions regarding the bilinguals' VOT:

1) Do Dutch-German bilingual children produce language-specific VOT in Dutch and in German and is more exposure to German associated with longer VOT in both languages?
2) Do Dutch-German bilingual children differ from monolingual children in their Dutch and German VOT production?
3) Is VOT associated with age in Dutch-German bilingual and monolingual preschoolers?

If the bilingual children are subject to CLI, their VOT productions should differ from those of monolinguals in at least one language. Given that the bilinguals' majority language is Dutch, an influence from Dutch to German is expected to be more prominent than the influence from German to Dutch. The SLM's Age of Acquisition Hypothesis (Flege, 1995) suggests that bilinguals acquire language-specific categories for 'voiceless' and 'voiced' plosives. By contrast, a prediction that follows from the SLM's Equivalence Classification Hypothesis (Flege, 1987, 1995) is that the 'voiceless' plosives of the two languages may be merged to one single category, and similarly, the 'voiced' plosives of the two languages may be merged into one category. Based on the A-Map model (McAllister Byun et al., 2016), it is expected that bilinguals may not acquire prevoicing in Dutch and aspiration in German at the same age as their monolingual peers. This is because bilingual children have accumulated less production experience with these articulatory and aerodynamically complicated sounds in their two languages relative to their monolingual peers. Similarly, bilingual children with more exposure to German, and therefore more heritage language experience, are predicted to be more successful in producing target-like VOT in German, and may consequently be less successful in producing target-like VOT in Dutch than bilingual children with less exposure to German. Finally, because anatomical and motor-control constraints may be decreasing between 3;7 and 6;0 years, older bilingual and monolingual children are expected to produce prevoicing and aspiration more reliably than younger children.

2 Method

2.1 Participants

Eighty-eight children between 3;6 and 6;0 years participated in this study: 29 Dutch-German bilinguals (M_{age}=4;7, range 3;7–5;11; 14 female), 30 Dutch monolinguals (M_{age}=4;9, range 3;6–6;0; 17 female) and 29 German monolinguals (M_{age}=4;8, range 3;6–6;0; 20 female)[3]. The groups did not differ significantly in age, $F(2,85)=0.5$, $p>.250$.

Of the initially tested 97 children, four bilinguals were excluded either due to exposure to a third language (N=3) or onset of bilingualism after the first year of life (N=1). Five monolinguals were excluded either due to exposure to foreign accented speakers (N=4) or inability to complete the task (N=1). Based on parental report, all children were typically developing and had no speech impairments or delays, and no auditory, cognitive or neurological impairments. Only bilinguals able to communicate in Dutch and German participated.

The children were recruited from the participant pools of the Baby Research Center Nijmegen and the University of Amsterdam, or via online and offline classifieds. The bilingual children were tested in different regions of the Netherlands (Gelderland (N=16), Amsterdam (N=9), Utrecht (N=2), Limburg (N=1), North Brabant (N=1)). All monolingual Dutch children were tested in Gelderland in the Central Eastern Netherlands. The monolingual German children were tested in Central Western Germany (N=27) and Northern Germany (N=2).

Twenty bilingual children had a German mother and a Dutch father, and six had a Dutch mother and a German father. Three children had two German parents, but were born in the Netherlands. Two of them were exposed to Dutch through native speakers from birth. The third child's first regular exposure to Dutch started at 0;6.

[3] The bilingual group contained one set of siblings and the German monolingual group contained three sets of siblings. Removing siblings from the data did not change the pattern of results.

Detailed assessments of language exposure based on the Bilingual Language Experience Calculator (BiLEC; Unsworth, 2013) revealed that the bilingual children had on average more exposure to Dutch (*M*=58%, range 22%–89%, *SD*=15) than to German (*M*=42%, range 11%–78%, *SD*=15) at the time of testing, $t(28)=2.89$, $p=.007$. These percentages were based on parents' responses to questions inquiring how much time children spent at daycare or school, and with various family members. Furthermore, the BiLEC asks which languages the relevant family members speak toward the child, allowing answers such as 70% Dutch and 30% German. An algorithm within the BiLEC then calculates a child's current amount of exposure (in percent) to a language.

Parents provided proficiency ratings for their child's ability to speak and understand each language on a scale from 0 (virtually no fluency; almost no understanding) to 5 (native fluency, native understanding). The bilinguals were assigned better speaking scores in Dutch (*M*=4.6, range 2–5, *SD*=0.8) than in German (*M*=3.3, range 1–5, *SD*=1.3), $t(28)=4.23$, $p<.001$. Similarly, their ability to understand Dutch (*M*=4.9, range 3–5, *SD*=0.4) was rated better than their ability to understand German (*M*=4.6, range 3–5, *SD*=0.6), $t(28)=2.29$, $p=.030$. According to self-report, the parents of the bilinguals had the highest education[4] (mothers: *M*=5.3, range 2–6; fathers: *M*=5.3, range 4–6), followed by the parents of the Dutch monolinguals (mothers: *M*=5, range 3–6; fathers: *M*=4.7, range 2–6) and the parents of the German monolinguals (mothers: *M*=4.7, range 2–5; fathers: *M*=3.3, range 2–5), $F(2,80)=17.73$, $p<.001$ for mothers and $F(2,80)=24.53$, $p<.001$ for fathers. Bonferroni post hoc tests revealed that only the mothers and fathers of the German monolinguals had significantly lower education than the mothers and fathers in the other two groups.

2.2 Materials and procedure

2.2.1 Target words

The investigated plosives were 'voiceless' /p/, /t/ and /k/ and 'voiced' /b/ and /d/. The 'voiced' dorsal plosive /g/ is not a native phoneme in Dutch, and is therefore not addressed in this study. For each of the five plosives, a total of six target words per language was selected from the Dutch version of the MacArthur-Bates

[4] 0=Kindergarten, 1=Elementary school, 2=10th grade, 3=High school, 4=Bachelor's degree, 5=Master's degree, 6=Doctorate.

Communicative Development Inventories (Zink & Lejaegere, 2002), and for German from the questionnaire on early child language development (Szagun, Stumper & Schramm, 2009) as well as from the parental questionnaire on early diagnosis of at-risk children (Grimm & Doil, 2000). A list of all target words can be found in Appendix 1.

All target words were picturable plosive-vowel-initial nouns. Due to restrictions in the availability of suitable target words, no match in vocalic contexts between Dutch and German target words could be achieved. We address this issue in Appendix 2 with descriptive statistics showing how the children's VOT differs by vocalic context. Appendix 2 is supplemented by an additional analysis supporting that the imbalance of the vocalic context in Dutch and German did not influence the results reported in this study.

2.2.2 Elicitation materials and procedures

To keep the children's attention during the test session, their speech was elicited in two different picture naming tasks. The pictures in both tasks were color photographs and color drawings[5]. Each production of the child entered the analysis. If a child did not say the target word, the experimenter gave the child hints without saying the word or the initial plosive. If the child still did not know the word after three elicitation attempts, the experimenter named the picture and continued to the next trial.

In the story task, the experimenter read a custom-made story to the child, which was set up in Microsoft Powerpoint and presented on a laptop computer. Within the story, the target words were replaced with pictures representing the words. Each picture occurred on a separate slide and the children were instructed to name each picture.

The picture naming game was designed as a lotto matching game. The child and the experimenter each had one lotto board in DIN A3 format. Half of the pictures were printed on the child's lotto board and the other half were printed on the lotto board of the experimenter. The same pictures were printed on individual 6 x 6 cm cards, which had to be matched to one of the two lotto boards. The child and the experimenter took turns in turning around one card at a time.

During the game, the experimenter instructed the child to name the pictures for a hand puppet. A koala hand puppet, which was introduced as being unfamiliar

[5] Six out of 30 pictures in the Dutch task and four out of 30 pictures in the German task were color drawings because they represented the words better than photographs.

with the words was used in the Dutch session, and an allegedly blind mole hand puppet was used in the German session.

The player who first collected all matching picture cards on her lotto board won the game. The order of the cards was randomized with two exceptions: to let the child win the game, the last card on the stack always belonged to the experimenter's lotto board. In addition, the second last card always belonged to the child's lotto board, which ensured that the game did not end before all pictures were named.

2.2.3 General procedure

Testing took place in a quiet room at the children's homes. As outlined below, data for all experiments reported in this dissertation were collected in two test sessions (one Dutch session and one German session). The bilinguals were tested by native speakers of the test language, and the two sessions were scheduled approximately two weeks apart. Half of the children completed the Dutch session before the German session, and the other half started with the German session.

At the beginning of the first session, parents signed informed consent. Afterwards, the parents completed a picture naming task for the study reported in Chapter 5. The parents then completed a language background questionnaire. The questionnaire for bilingual children was based on the BiLEC (Unsworth, 2013), and the monolingual version was custom-made and screened for potential exposure to additional languages and foreign accents.

The first task the children completed was the picture naming story. After the story task, the children completed a speech perception experiment for the study reported in Chapter 3. The final task of the session was the picture naming game.

Throughout the session, children were rewarded with stickers. Breaks were taken in between all tasks. Additional breaks were taken as needed. At the end of each session, the children were compensated with their choice between €10 or a book.

2.3 Recordings and VOT measurements

Recordings were made with an Olympus Linear PCM Recorder LS-10 with uncompressed 24bit/96kHz recording capability. The first author measured VOT of all children in Praat (Boersma & Weenink, 2015) taking into account waveforms and spectrograms viewed at 0–5000 Hz. Burst onset was defined as the onset of abrupt energy release. If there was more than one release burst, VOT was measured from the first visible release burst (Mayr & Siddika, 2016). Onset of voicing was defined as the first periodic component of the waveform and was measured at the preceding zero-crossing (Francis, Ciocca & Man Ching Yu, 2003). When the amplitude increase of prevoicing was gradual, voicing onset measurements were based on visual characteristics. Figure 2 provides examples of VOT measurements in the prevoicing, short lag, and aspiration ranges, respectively. Three additional phonetically trained coders measured 25% of the data. Inter-coder reliability indicated 98% agreement. For 'voiceless' plosives, measurements were considered in agreement when they differed in less than 10 ms (Fabiano-Smith & Bunta, 2012). Measurements of 'voiced' plosives were considered in agreement when both coders rated VOT as either prevoiced or devoiced. Across groups and plosives, 11% of the tokens were excluded from the analyses because they could not be unambiguously measured, for example, due to coarticulation, sound overlap, creaky voice, or whispering.

2.4 Statistical analyses

Mixed effects models were performed in R (R Core Team, 2013). An α-level of .05 was adopted throughout. For the 'voiceless' plosives /p/, /t/ and /k/, mixed effects linear regression was performed with VOT as the dependent variable. Initial data screening revealed a bimodal distribution of VOT in the 'voiced' plosives in 59/60 children in Dutch and 46/59 children in German. As presence versus absence of prevoicing rather than duration of prevoicing plays a crucial role in Dutch (van Alphen & McQueen, 2006), VOT was converted into a categorical variable with the levels 'prevoiced' for negative VOT and 'devoiced' for positive VOT. This categorical dependent variable entered a mixed effects logistic regression.

Several independent variables (IVs) were used in the models. Language (Dutch, German) was the IV of main interest in within-group analyses that compared the bilinguals' two languages, and also in between-group analyses involving the two monolingual groups. Language Background (monolingual, bilingual) was the IV of main interest in the between-group comparisons of bilinguals and monolinguals that were conducted separately for Dutch and German. The IV Age (in months) was

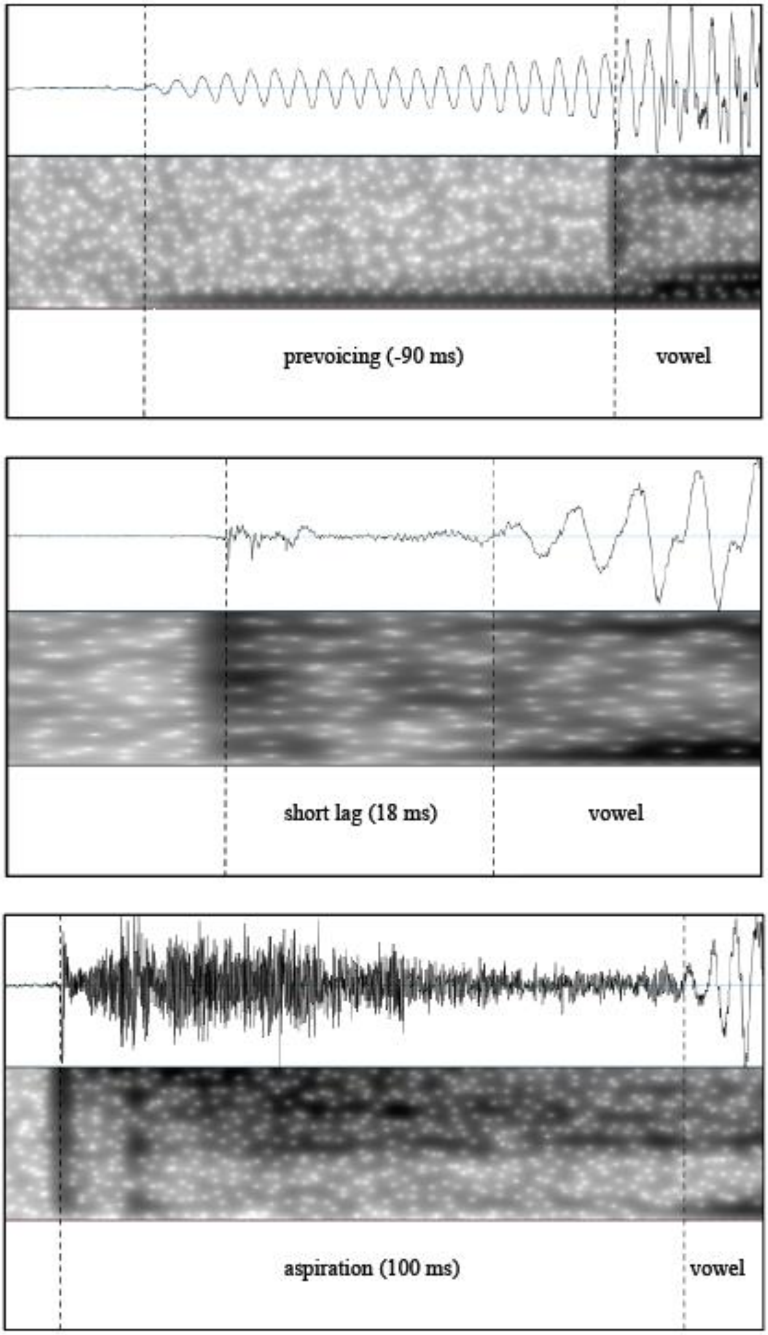


Figure 2. Acoustic landmarks from top to bottom: prevoicing, short lag, and aspiration.

included in all analyses, and Percent of Exposure to German[6] was only included in the within-group analyses on the bilinguals. These latter two IVs were centered around zero for each analysis.

Three additional IVs were included in the models: Elicitation Task of the item, Place of Articulation of the plosive, and Word Length ('voiceless' plosives only) of the item. These additional IVs were merely included to account for variance in the data, but did not contribute to the main results reported here. Simple effects of these IVs can be found in Appendix 3.

Table 3 provides an overview of the model specifications including fixed effects, interaction terms, random effects, intercepts, and random slopes for each group comparison. All models include interaction terms between the IV of main interest and all secondary IVs. Significant interactions are reported below, and information on post-hoc analyses is provided in Appendix 4.

3 Results

This section starts with the descriptive statistics before we turn to the statistical effects of Language and Language Background on VOT, taking into account the children's age, and in case of the language comparison within the bilinguals, their exposure to German.

For 'voiceless' plosives, monolingual Dutch children produced the shortest and German monolingual children the longest average VOT. The bilinguals' VOT was intermediate to the two monolingual groups. The bilinguals further produced shorter VOT in Dutch than in German (see Table 4 & Figure 3).

For 'voiced' plosives, monolingual Dutch children produced the highest and German monolingual children the lowest percentage of prevoiced plosives. Bilinguals fell in between the monolinguals, with only a slightly higher percentage of prevoicing in Dutch than in German (see Table 5 and Figure 4). These percentages reflect the behavior of the vast majority of children, who prevoiced part of their 'voiced' plosives. Only 13 children (one bilingual speaking Dutch, three bilinguals speaking German, and nine German monolinguals) never produced prevoicing. Conversely, only one child (a bilingual speaking German) produced all 'voiced' plosives with prevoicing. In Dutch, only six monolingual children and three bilingual children fell within the adult-like 75–100% range of prevoicing.

[6] Percent of exposure to German is inversely proportional to percent of exposure to Dutch.

Table 3. Model specifications.

Groups	Analysis	Fixed effects	Interactions	Random effects & intercept	Random slopes
Monolingual Dutch vs. monolingual German	‘voiceless’	Language Age Task PoA-LC PoA-CD WordLength	Language*Age Language*Task Language*PoA-LC Language*PoA-CD Language*WordLength	Child	Task PoA-LC PoA-CD WordLength
				Item	Task
	‘voiced’	Language Age Task PoA	Language*Age Language*Task Language*PoA	Child	Task PoA
				Item	Task
Bilingual Dutch vs. bilingual German	‘voiceless’	Language Age %ExposureG Task PoA-LC PoA-CD WordLength	Language*Age Language*%ExposureG Language*Task Language*PoA-LC Language*PoA-CD Language*WordLength	Child	Language Task PoA-LC PoA-CD WordLength
				Item	Task
	‘voiced’	Language Age %ExposureG Task PoA	Language*Age Language*%ExposureG Language*Task Language*PoA	Child	Language Task PoA
				Item	Task
Bilinguals vs. monolinguals	‘voiceless’	LangBackgr. Age Task PoA-LC PoA-CD WordLength	LangBackgr.*Age LangBackgr.*Task LangBackgr.*PoA-LC LangBackgr.*PoA-CD LangBackgr.*WordLength	Child	Task PoA-LC PoA-CD WordLength
				Item	LangBackgr. Task
	‘voiced’	LangBackgr. Age Task PoA	LangBackgr.*Age LangBackgr.*Task LangBackgr.*PoA	Child	Task PoA
				Item	LangBackgr. Task

%ExposureG= Percent of Exposure to German
LangBackgr.=Language Background
PoA-LC=Place of Articulation: Labial vs. Coronal
PoA-CD=Place of Articulation: Coronal vs. Dorsal

Table 4. 'Voiceless' plosives: mean VOT values (ms) by language and language background over children.

		Dutch		German	
		Monolinguals	Bilinguals	Bilinguals	Monolinguals
/p/	***M***	**15**	**20**	**31**	**52**
	SD	16	21	30	27
	Total #	371	353	431	391
/t/	***M***	**23**	**28**	**56**	**77**
	SD	19	26	42	36
	Total #	386	373	428	401
/k/	***M***	**32**	**40**	**57**	**75**
	SD	25	29	41	32
	Total #	379	383	428	391
	Group Total #	1136	1109	1287	1183

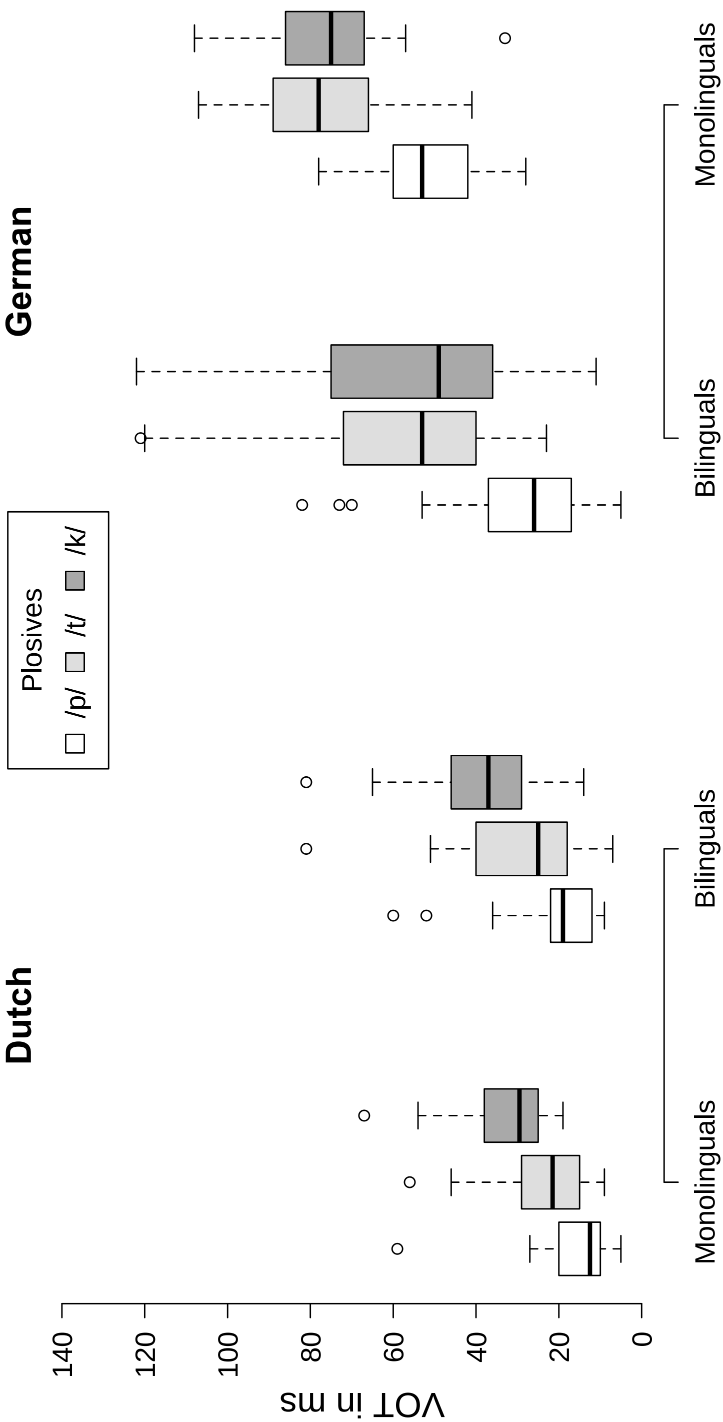


Figure 3. 'Voiceless' plosives: VOT by language and language background over children.

Table 5. ‘Voiced’ plosives: mean percentage of prevoiced plosives by language and language background over children.

		Dutch		German	
		Monolinguals	Bilinguals	Bilinguals	Monolinguals
	M %	**56**	**33**	**30**	**12**
/b/	*SD*	27	27	27	13
	Total #	177/316	112/340	112/370	48/381
	M %	**43**	**28**	**21**	**4**
/d/	*SD*	25	29	26	8
	Total #	155/349	97/344	83/412	15/429
	Range prevoiced	9–96%	0–95%	0–100%	0–41%
	Total # /b/ + /d/	332/665	209/684	195/782	63/810

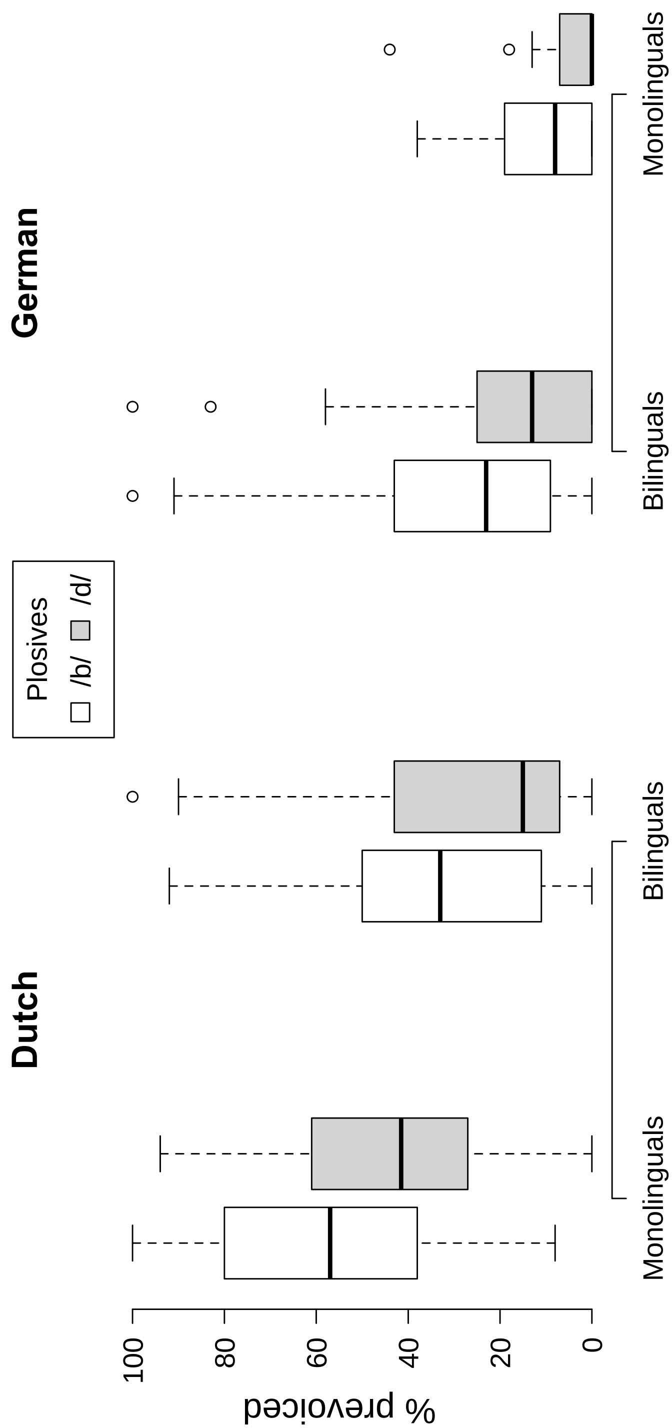


Figure 4. 'Voiced' plosives: percent prevoiced by language background over children.

The devoiced 'voiced' plosives fell on average within the short lag VOT range. All groups produced devoiced /b/ with VOT of around 10 ms. For devoiced /d/, the Dutch monolinguals and the bilinguals in both languages produced VOT of around 20 ms. The German monolinguals produced shorter VOT with a mean of 13 ms (see Table 6). All groups produced shorter VOT for devoiced 'voiced' plosives than for 'voiceless' plosives, but this difference is very small in the group of Dutch monolingual children (see Tables 4 & 6). Figure 5 shows the distribution of VOT across all 'voiced' plosives by group and language.

Table 6. 'Voiced' plosives: mean VOT values (ms) of devoiced 'voiced' plosives by language and language background over children.

		Dutch		German	
		Monolinguals	Bilinguals	Bilinguals	Monolinguals
	M	**11**	**13**	**11**	**10**
/b/	*SD*	7	5	6	4
	Total #	139	229	258	333
	M	**20**	**21**	**18**	**13**
/d/	*SD*	10	7	5	4
	Total #	194	247	329	414
	Group Total#	333	476	587	747

Four sets of mixed effects analyses were performed, and Table 7 summarizes the results. Two initial analyses confirmed that monolingual Dutch children and monolingual German children differ in their VOT production. As expected, monolingual Dutch children produced 'voiceless' plosives with overall shorter VOT than monolingual German children (β=28.96, *SE*=2.95, *t*=9.82, *p*<.001). Interactions between Language and Place of Articulation (labial vs. coronal; β=-7.28, *SE*=3.46, *t*=-2.10, *p*=.036) as well Language and Word Length (β=4.33, *SE*=1.36, *t*=3.18, *p*=.002) indicated that the German monolingual children produced shorter VOT in labial /p/ than in coronal /t/ (β=-21.34, *SE*=5.89, *t*=-3.63, *p*<.001) and longer VOT in monosyllabic than in disyllabic words (β=8.79, *SE*=2.37, *t*=3.71, *p*<.001), but neither effect was observed in the monolingual Dutch children (β=-6.33, *SE*=3.60, *t*=-1.76, *p*=.079 and β=0.09, *SE*=1.31, *t*=.06, *p*>.250, respectively). Monolingual Dutch children produced a higher percentage of 'voiced' plosives with prevoicing than monolingual German children (β=1.56, *SE*=0.19, *z*=8.07, *p*<.001). An

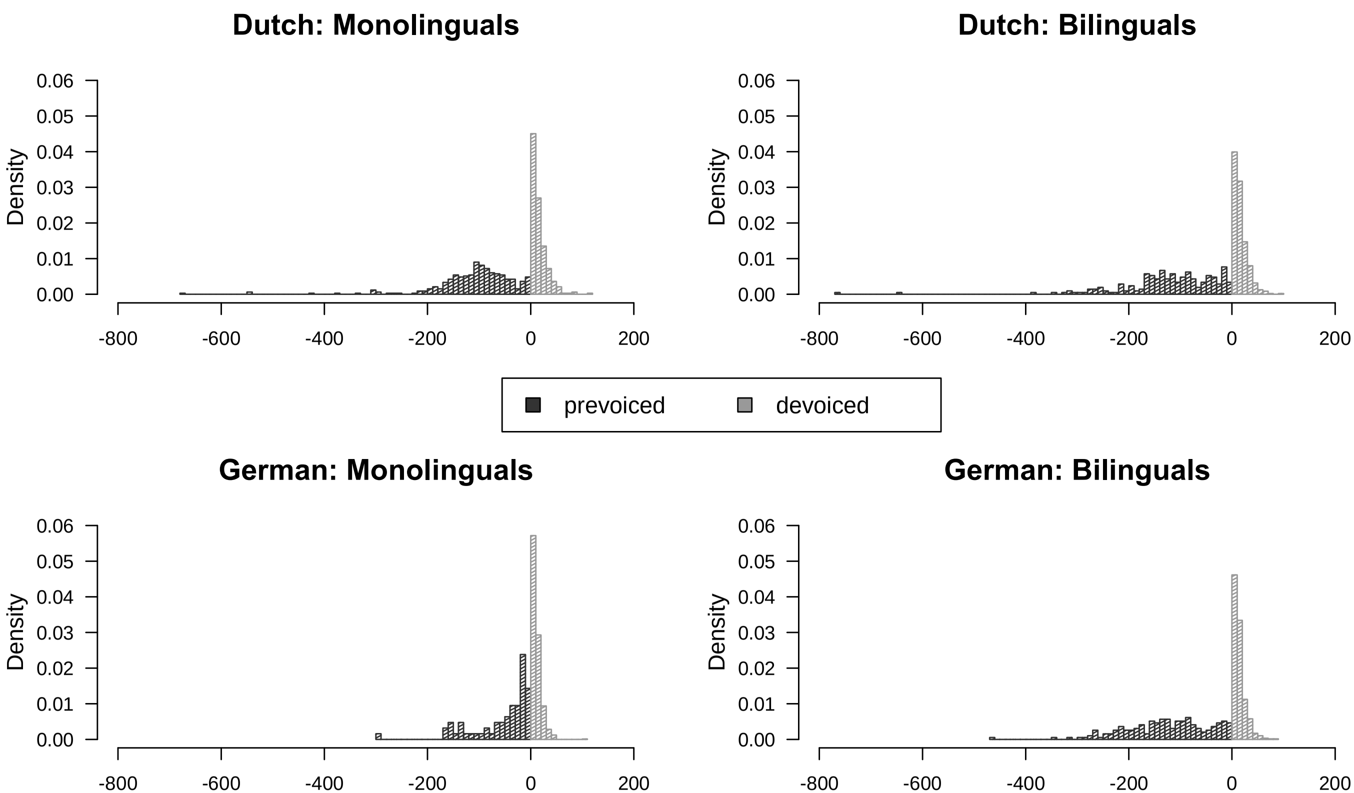


Figure 5. VOT distribution of ‘voiced’ plosives by language and language background.

interaction between Language and Place of Articulation (β=-0.21, *SE*=0.11, *z*=-2.03, *p*=.042) indicated that both groups prevoiced labial /b/ more frequently than coronal /d/, but the magnitude of the effect was larger in the German monolinguals (β=-1.06, *SE*=0.23, *z*=-4.52, *p*<.001) than in the Dutch monolinguals (β=-0.32, *SE*=0.12, *z*=-2.80, *p*=.005). The observed differences between monolingual Dutch and German children are in line with the documented difference between Dutch and German plosives in adults' speech.

The next analyses tested whether Dutch-German bilingual children produce language-specific VOT in Dutch and in German and whether their relative heritage language exposure is associated with their VOT. Dutch-German bilingual children produced 'voiceless' plosives with longer VOT in German than in Dutch (β=14.43, *SE*=3.54, *t*=4.08, *p*<.001). An interaction between Language and Percent of Exposure to German (β=0.26, *SE*=0.09, *t*=2.85, *p*=.004) revealed that more exposure to German is associated with longer, and therefore more target-like, VOT in German (β=0.52, *SE*=0.24, *t*=2.17, *p*=.030), while it had no detectable effect on the bilinguals' Dutch VOT (β=0.15, *SE*=0.14, *t*=1.09, *p*>.250) as visualized in Figure 6. Similarly, an interaction between Language and Task (β=1.07, *SE*=0.49, *t*=2.21, *p*=.027) indicated that the bilinguals produced longer VOT in the game task than in the story task in Dutch (β=-3.14, *SE*=0.92, *t*=-3.42, *p*<.001), but not in German (β=-1.11, *SE*=1.25, *t*=-0.89, *p*>.250). The percentage of 'voiced' plosives produced with prevoicing was similar in the bilinguals' Dutch and German (β=0.26, *SE*=0.17, *z*=1.5, *p*=.134), and it was not significantly affected by Percent of Exposure to German (β=0.02, *SE*=0.02, *z*=1.22, *p*=.223).

The following analyses tested whether Dutch-German bilingual children produce VOT differently than their monolingual peers. For Dutch 'voiceless' plosives, no significant VOT differences were observed between Dutch-German bilingual children and monolingual Dutch children (β=2.86, *SE*=1.92, *t*=1.50, *p*=.134). However, the bilinguals produced a lower percentage of prevoiced 'voiced' plosives in Dutch than their monolingual peers (β=0.51, *SE*=0.21, *z*=2.40, *p*=.016).

In German, the Dutch-German bilingual children produced 'voiceless' plosives with overall shorter, and therefore more Dutch-like, VOT than monolingual German children (β=-10.2, *SE*=3.12, *t*=-3.27, *p*=.001). Similarly, the bilingual children prevoiced a higher percentage of 'voiced' plosives in German than their monolingual peers (β=-0.94, *SE*=0.25, *z*=-3.72, *p*<.001). An interaction between Language Background and Place of Articulation (β=0.27, *SE*=0.13, *z*=2.01, *p*=.044) indicated that both groups prevoiced labial /b/ more frequently than coronal /d/, but the magnitude of the effect was larger in the monolinguals (β=-1.06, *SE*=0.23, *z*=-4.52, *p*<.001) than in the bilinguals (β=-0.41, *SE*=0.15, *z*=-2.79, *p*=.005).

No effects of Age and no interactions between Language and Age or Language Background and Age were observed either in 'voiceless' or in 'voiced' plosives in any of the analyses (monolingual Dutch vs. monolingual German: 'voiceless': β=0.17, *SE*=0.15, *t*=1.2, *p*=.230 & 'voiced': β=-0.02, *SE*=0.02, *z*=-1.14, *p*>.250; bilingual Dutch vs. bilingual German: 'voiceless': β=-0.23, *SE*=0.29, *t*=-0.80, *p*>.250 & 'voiced': β=-0.02, *SE*=0.03, *z*=-0.72, *p*>.250; bilingual Dutch vs. monolingual Dutch: 'voiceless': β=0.09, *SE*=0.14, *t*=0.62, *p*>.250 & 'voiced': β=-0.03, *SE*=0.02, *z*=-1.57, *p*=.116; bilingual German vs. monolingual German: 'voiceless': β=-0.15, *SE*=0.25, *t*=-0.58, *p*>.250 & 'voiced': β=-0.03, *SE*=0.02, *z*=-1.05, *p*>.250).

2

4 Discussion

This study examined bilingual preschoolers' VOT development in their majority language Dutch and their heritage language German, in comparison to age-matched monolingual peers. In the following, the findings are summarized and explained in terms of CLI and language exposure. We specifically discuss whether these two more general constructs can be captured by the A-Map model (McAllister Byun et al., 2016) and the Speech Learning Model's Age of Acquisition and Equivalence Classification Hypotheses (Flege, 1995). We first discuss the children's production of 'voiceless' plosives and then turn to the production of 'voiced' plosives.

In sum, the bilingual and monolingual children's production of VOT in 'voiceless' plosives revealed three main findings, and an initial analysis confirmed the expected differences between Dutch and German monolingual preschoolers. The bilingual children's productions provide evidence for language-differentiation between their Dutch and German phonetic systems, and furthermore reveal an effect of language exposure on VOT in the heritage language German, but not on the majority language Dutch (Research Question 1). Moreover, the bilinguals produced VOT differently from their monolingual peers in the heritage language German, but not in the majority language Dutch (Research Question 2). Finally, we did not observe an age-effect on VOT (Research Question 3).

Monolingual Dutch children produced 'voiceless' plosives with short lag VOT whereas monolingual German children produced aspiration, which is in line with Dutch and German adults' VOT production, respectively (Deighton-van Witsen, 1976; Fischer-Jørgensen, 1976; Haag, 1979; Jessen, 1998; Lisker &

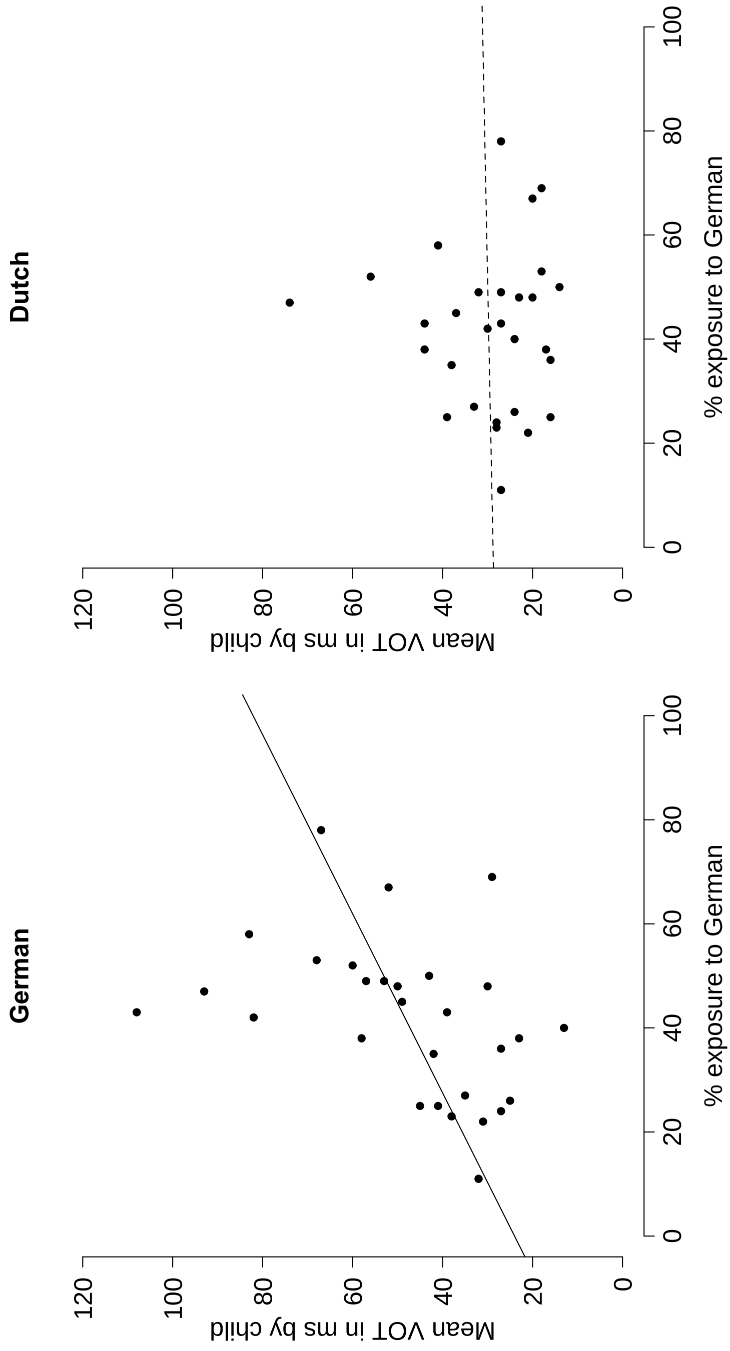


Figure 6. VOT of 'voiceless' plosives and exposure to German over children.

Table 7. Summary: effects of Language and Language Background on VOT.

		Monolingual Dutch group vs. monolingual German group	**RQ 1 Bilingual group Dutch vs. German**	**RQ 2 Bilingual group in Dutch vs. monolingual Dutch group**	**Bilingual group in German vs. monolingual German group**
Language	'voiceless'	Longer VOT in German ***	Longer VOT in German ***	--	
	'voiced'	Higher % of prevoicing in Dutch ***	n.s.	--	--
Language Background	'voiceless'	--	--	n.s.	Bilinguals: shorter VOT ***
	'voiced'	--	--	Bilinguals: lower % of prevoicing *	Bilinguals: higher % of prevoicing ***
% Exposure to German	'voiceless'	--	More exposure to German → longer VOT in German *	--	--
	'voiced'	--	n.s.	--	--
RQ 3: Age	'voiceless'	n.s.	n.s.	n.s.	n.s.
	'voiced'	n.s.	n.s.	n.s.	n.s.

*** $p<.001$; ** $p<.01$; * $p<.05$; n.s. $p>.05$

Abramson, 1964; Neuhauser, 2011). Equivalent to Dutch and German monolingual children, the bilinguals produced longer VOT in German than in Dutch, suggesting that bilingual children have separate phonological categories for Dutch and German 'voiceless' plosives. This finding is in line with the SLM's Age of Acquisition Hypothesis, which suggests that early bilingual acquisition promotes language-specific category formation. Importantly, those bilingual children with more exposure to German produced longer, and therefore more German-like VOT in German, but more exposure to German did not detectably influence their Dutch VOT. Previous research on Welsh-English bilinguals similarly revealed effects of language exposure on the minority language, but not on the majority language (Gathercole & Thomas, 2009). These results indicate that more heritage language exposure is beneficial to the development of the heritage language, but not at the cost of the counterpart category in the majority language. As needs to be confirmed by future research, the bilingual children's Dutch VOT is presumably not perceived as foreign-accented, even when exposure to the heritage language German is high (Flege, 1984; Major, 1987; Riney & Takagi, 1999; Sancier & Fowler, 1997; Schoonmaker-Gates, 2015).

Despite the bilinguals' production of aspiration in German, they produced 'voiceless' plosives with shorter VOT than monolingual German children. Differences between bilinguals and monolinguals in absolute VOT duration in German may be related to CLI and differences in exposure to German.

CLI from Dutch to German may cause the bilinguals' shorter VOT durations in German, suggesting that their separate 'voiceless' categories for Dutch and German interact. Such CLI has often been reported for bilingual children across different languages (Fabiano & Goldstein, 2005; Fabiano-Smith & Bunta, 2012; Kehoe, 2002; Kehoe et al., 2004; Lleó & Kehoe, 2002; Mayr & Siddika, 2016).

Language exposure was a crucial factor impacting on the German VOT in the bilingual group, suggesting that differences in language exposure between bilingual and monolingual children can similarly account for differences in VOT duration between the two groups. The A-Map model captures these differences in language exposure within the group of bilinguals and also between the bilinguals and monolinguals. All children in this study are clearly beyond the critical age of 2;0 at which monolingual children start producing aspiration (Kager et al., 2007; Macken & Barton, 1980a), but the bilinguals' exposure to German is limited to 42% of their waking hours on average. Compared to the monolingual A-Map, the bilingual A-Map is therefore based on less experience in the production of aspiration, which can explain why the bilinguals produced more variable and overall shorter aspiration than monolingual children.

The specific A-Maps of bilingual children can further differ between children as a result of individual differences in language experience. More experience with German could increase the urge of bilingual children to reproduce the adult aspiration target accurately, as well as provide them with more practice to reach that target precisely. However, this experience and precision in aspirating in the heritage language German does not result in the children abandoning the fully accurate and precise short lag VOT of ‘voiceless’ plosives in the majority language Dutch. Individual differences in language experience suggest that the Dutch and German ‘voiceless’ categories may in fact be separate and autonomous. Note, however, that a lack of surfacing CLI cannot preclude the existence of CLI.

Specific analyses on the bilingual children’s production of ‘voiced’ plosives revealed three main findings, and confirmed the expected production differences between monolingual Dutch and German children. First, we did not observe language-differentiation between the bilinguals’ Dutch and German ‘voiced’ plosives, and a child’s language exposure was not detectably associated with her production of ‘voiced’ plosives (Research Question 1). Second, the bilinguals’ productions of ‘voiced’ plosives differed from monolinguals’ productions in the heritage language German and also in the majority language Dutch (Research Question 2). Third, no age-effect on the percentage of prevoiced ‘voiced’ plosives was observed (Research Question 3).

Monolingual German children primarily produced devoiced ‘voiced’ plosives and only prevoiced about 10% of them. These findings are in line with previous research on German toddlers (Kehoe et al., 2004). The monolingual children’s German productions fall within adult ranges in the distribution of prevoiced and devoiced ‘voiced’ plosives (Fischer-Jørgensen, 1976; Haag, 1979; Jessen, 1998; Neuhauser, 2011; Stoehr, Benders, van Hell & Fikkert, 2017 / Chapter 5).

Monolingual Dutch children prevoiced about 50% of their ‘voiced’ plosives and devoiced the remaining 50%. This percentage is below the adult-target of 75% to 100% of prevoiced ‘voiced’ plosives in Dutch (Stoehr et al., 2017 / Chapter 5; van Alphen & Smits, 2004). Previous research on different languages similarly reported devoicing of target prevoiced plosives, possibly lasting into the early school years, and suggests that prevoicing is inherently difficult to produce (Allen, 1985; Bortolini et al., 1995; Kager et al., 2007; Kewley-Port & Preston, 1974; Khattab, 2000; Macken & Barton, 1980b; MacLeod, 2017). The A-Map model can explain the high within-child variation in prevoicing and devoicing of ‘voiced’ plosives by the monolingual Dutch children as a result of the competing pressures to accurately reproduce the adult-target (i.e., prevoicing) and to achieve a precise production (i.e., short lag) with a still-developing anatomy and motor control. The high variability across the monolingual Dutch children can be accounted for in terms of different rankings of these competing constraints.

Bilingual children prevoiced to a similar extent in Dutch (30%) and German (25%) and their percentages of prevoiced plosives fall in between the two monolingual groups. According to the A-Map model extended to bilingualism, the bilingual children's low percentage of prevoiced 'voiced' plosives in Dutch suggests that they are more affected by the constraint to achieve a precise production (i.e., short lag) than their monolingual peers. Possibly, less exposure to the 'prevoiced' adult-target makes the urge to reproduce prevoicing accurately relatively less impactful. The ranking of the constraints to achieve a precise production and to accurately match the adult-target may change with increasing language experience.

However, within the group of bilinguals, neither age nor their wide range of exposure to Dutch (22–89% of the children's waking hours) was detectably associated with the bilinguals' production of prevoicing in Dutch or German. This also renders it unlikely that differences in exposure to Dutch between bilinguals and monolinguals can account for the groups' different percentages of prevoicing. Hence, the A-Map model cannot entirely account for the bilinguals' differential production of 'voiced' plosives.

Instead, bidirectional CLI can explain the bilinguals' production of 'voiced' plosives. In this case, CLI may be captured through equivalence classification or acceleration. The SLM's Equivalence Classification Hypothesis predicts that CLI results in the formation of a single category for two perceptually close sounds from two languages. Accordingly, Dutch-German bilingual children appear to have only one 'voiced' category for Dutch and German. The bilinguals may be in the process of approaching the prevoiced Dutch adult-target with this merged 'voiced' category, as they produce prevoicing in German, which is articulatory and aerodynamically complex and unlikely to result from any default behavior (Kewley-Port & Preston, 1974). This merger would effectively take the German 'voiced' category out of the short lag VOT range and eliminate the double phonological function of the short lag VOT range, which otherwise corresponds to 'voiceless' in Dutch and to 'voiced' in German. The hypothesized merger may eventually match the target Dutch phonology, in which prevoicing is crucial for the realization of the voicing opposition without violating the target German phonology, in which prevoicing occurs as free variation (Fischer-Jørgensen, 1976; Hamann & Seinhorst, 2016; Jessen, 1998; Stock, 1971; Stoehr et al., 2017 / Chapter 5).

However, the present data are also compatible with the hypothesis that the bilinguals have two separate 'voiced' categories for Dutch and German that develop indistinguishably at the current developmental stage. In this case, CLI occurs as acceleration from Dutch to German, and can be explained by the A-Map model. Similar acceleration effects in the domain of phonology have previously been reported in bilingual children of different language backgrounds (Grech & Dodd, 2008; Mayr et al., 2015; Tamburelli et al., 2015). The bilinguals prevoiced more frequently in German (25% of all 'voiced' plosives) than monolingual German children (8% of all 'voiced' plosives; Kehoe et al., 2004). German adults prevoice on average up to 50% of 'voiced' plosives, which means that the bilingual children are in fact closer to the adult-target than their monolingual peers (Fischer-Jørgensen, 1976; Hamann & Seinhorst, 2016; Jessen, 1998; Stock, 1971; Stoehr et al., 2017 / Chapter 5). The bilinguals' exposure to Dutch leads to more exposure to prevoicing, and to more experience producing it. In line with the A-Map model, bilinguals accumulate prevoicing experience in Dutch, and their episodic memory therefore encompasses more traces of the articulator movements associated with prevoicing. This production experience may accelerate the bilinguals' acquisition of this typically late-acquired structure in German. Assuming acceleration in German, the bilingual children's percentage of prevoiced 'voiced' plosives should increase in German until they reach similar variation between prevoicing and short lag VOT as observed in German-speaking adults. The Dutch category should then keep developing to the adult-target of 75%–100% of prevoicing. Speech perception or longitudinal speech production research is needed to identify whether CLI in bilingual children's production of 'voiced' plosives occurs as equivalence classification or acceleration.

5 Conclusions

This study contributed new insights into the role of heritage language exposure in bilingual children's VOT development. The results extend findings of previous small-scale studies through evidence that inherently difficult prevoicing is not only prone to differential acquisition in a heritage language, as previously reported, but also in a majority language. The bilinguals' similar production of prevoicing in both languages and the observed differences between bilinguals and monolinguals seem to be unrelated to variation in language exposure or age, and may instead result from

CLI. Moreover, aspiration can be prone to differential acquisition in a heritage language, especially when the exposure to the heritage language is low. Despite differences from monolingual VOT development, the bilinguals nevertheless seem to have acquired two separate and autonomous categories for Dutch and German ‘voiceless’ plosives. Importantly, this study revealed a positive effect of more heritage language exposure on the production of ‘voiceless’ plosives: bilingual children with more heritage language exposure produced more target-like VOT in the heritage language, but not at the cost of the majority language. What surfaces as CLI from Dutch to German in ‘voiceless’ plosives can be explained by language exposure alone. This novel evidence suggests that more exposure to the heritage language is associated with better-separated language-specific voicing systems.

References

Abramson, A. S., & Lisker, L. (1973). Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics*, *1*, 1–8.

Allen, G. D. (1985). How the young French child avoids the pre-voicing problem for wordinitial voiced stops. *Journal of Child Language*, *12*, 37–46.

Armon-Lotem, S., & Ohana, O. (2017). A CDI study of bilingual English-Hebrew children – frequency of exposure as a major source of variation. *International Journal of Bilingual Education and Bilingualism, 20*, 201–217.

Barton, D., & Macken, M. (1980). An instrumental analysis of the voicing contrast in word initial stops in the speech of four-year-old English-speaking children. *Language and Speech, 23*, 159–169.

Beckman, J., Helgason, P., McMurray, B., & Ringen, C. (2011). Rate effects on Swedish VOT: Evidence for phonological overspecification. *Journal of Phonetics, 39*, 39–49.

Boersma, P., & Weenink, D. (2015). *Praat: Doing phonetics by computer. Version 5.4.04*. Retrieved from http://www.praat.org/

Bortolini, U., Zmarich, C., Fior, R., & Bonifacio, S. (1995). Word-initial voicing in the productions of stops in normal and preterm Italian infants. *International Journal of Pediatric Otorhinolaryngology*, *31*, 191–206.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics, 27*, 207–229.

Davis, K. (1995). Phonetic and phonological contrasts in the acquisition of voicing: Voice onset time production in Hindi and English. *Journal of Child Language, 22*, 275–305.

Deighton-van Witsen, H. V. (1976). Preliminary investigation concerning differences between voice onset time values of speakers of standard Dutch and of one of the dialects spoken in the province of Groningen (second version). In Institute of Phonetic Sciences (Ed.), *Proceedings of the Institute of Phonetic Sciences, 4* (pp. 113–133). Retrieved from http://www.fon.hum.uva.nl/archive/1976/Proc4Deighton.pdf

Deuchar, M., & Clark, A. (1996). Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics, 24*, 351–365.

Eilers, R. E., Oller, D. K., & Benito-Garcia, C. R. (1984). The acquisition of voicing contrasts in Spanish and English learning infants and children: A longitudinal study. *Journal of Child Language, 11*, 313–336.

Fabiano, L., & Goldstein, B. (2005). Phonological transfer in bilingual Spanish-English speaking children. *Journal of Multilingual Communication Disorders, 3,* 56–63.

Fabiano-Smith, L., & Bunta, F. (2012). Voice onset time of voiceless bilabial and velar stops in 3-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics & Phonetics, 26*, 148–163.

Fabiano-Smith, L., & Goldstein, B. (2010). Phonological acquisition in bilingual Spanish-English speaking children. *Journal of Speech, Language, and Hearing Research*, *53*, 1–19.

Fischer-Jørgensen, E. (1954). Acoustic analysis of stop consonants. *Miscellaneo Phonetica, 2,* 42–59.

Fischer-Jørgensen, E. (1976). Some data on North German stops and affricates. *Annual Report of the Institute of Phonetics of the University of Copenhagen, 10*, 149–200.

Flege, J. E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, *76*, 692–707.

Flege, J. E. (1987). The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of Equivalence Classification. *Journal of Phonetics, 15*, 47–65.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America, 89*, 395–411.

Flege, J. E. (1995). Second language speech learning. Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Flege, J. E., Frieda, E. M., Walley, A. C., & Randazza, L. A. (1998). Lexical factors and segmental accuracy in second language speech production. *Studies in Second Language Acquisition, 20*, 155–187.

Flege, J. E., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech, 24*, 125–146.

Francis, A. L., Ciocca, V., & Man Ching Yu, J. (2013). Accuracy and variability of acoustic measures of voicing onset. *Journal of the Acoustical Society of America, 113*, 1025–1032.

Gathercole, V. M., & Hoff, E. (2007). Input and the acquisition of language: Three questions. In E. Hoff & M. Shatz (Eds.), *Handbook of language development* (pp. 107–127). Oxford: Blackwell.

Gathercole, V. M., & Thomas, E. M. (2005). Minority language survival: Input factors influencing the acquisition of Welsh. In J. Cohen, K. McAlister, K. Rolstad & J. MacSwan (Eds.), *Proceedings of the 4th International Symposium on Bilingualism* (pp. 852–874). Somerville, MA: Cascadilla Press.

Gathercole, V. M., & Thomas, E. M. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition, 12*, 213–237.

Gildersleeve-Neumann, C., & Wright, K. (2010). English speech acquisition in 3- to 5-year-old children learning Russian and English. *Language, Speech, and Hearing Services in Schools, 41*, 429–444.

Grech, H., & Dodd, B. (2008). Phonological acquisition in Malta: A bilingual language learning context. *International Journal of Bilingualism, 12*, 155–171.

Grimm, H., & Doil, H. (2000). *ELFRA: Elternfragebögen für die Früherkennung von Risikokindern, Handanweisung*. Göttingen: Hogrefe.

Haag, W. K. (1979). An articulatory experiment on voice onset time in German stop consonants. *Phonetica, 36*, 169–181.

Hamann, S., & Seinhorst, K. (2016). Prevoicing in Standard German plosives: Implications for phonological representations? Presented at the Thirteenth Old World Conference in Phonology, Budapest.

Heselwood, B., & McChrystal, L. (2000). Gender, accent features and voicing in Panjabi-English bilingual children. In D. Nelson & P. Foulkes (Eds.), *Leeds Working Papers in Linguistics and Phonetics, 8* (pp. 45–70). Retrieved from http://www.leeds.ac.uk/arts/info/125154/leeds_working_papers_in_linguistics_and_phonetics/1939/volume_8_2000

Jessen, M. (1998). *Phonetics and phonology of tense and lax obstruents in German*. Amsterdam: Benjamins.

Johnson, C., & Wilson, I. (2002). Phonetic evidence for early language differentiation: Research issues and some preliminary data. *International Journal of Bilingualism, 6*, 271–289.

Kager, R., van der Feest, S., Fikkert, P., Kerkhoff, A., & Zamuner, T. (2007). Representations of [voice]: Evidence from acquisition. In J. van de Weijer & E. J. van der Torre (Eds.), *Voicing in Dutch: (De)voicing – phonology, phonetics, and psycholinguistics* (pp. 41–80). Amsterdam: Benjamins.

Kehoe, M. (2002). Developing vowel systems as a window to bilingual phonology. *International Journal of Bilingualism, 6*, 315–334.

Kehoe, M. (2015). Cross-linguistic interaction: A retrospective and prospective view. In E. Babatsouli & D. Ingram (Eds.), *Proceedings of the International*

Symposium on Monolingual and Bilingual Speech 2015 (pp. 141–167). Retrieved from http://ismbs.eu/data/documents/Proceedings-ISMBS-2015.pdf

Kehoe, M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition, 7*, 71–88.

Kellerman, E., & Sharwood Smith, M. (1986). *Cross-linguistic influence in second language acquisition*. Elmsford, NY: Pergamon.

Kessinger, R. H., & Blumstein, S. E. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics, 25*, 143–168.

Kewley-Port, D., & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics, 2*, 194–210.

Khattab, G. (2000). VOT in English and Arabic bilingual and monolingual children. In D. Nelson & P. Foulkes (Eds.), *Leeds Working Papers in Linguistics and Phonetics, 8* (pp. 95–122). Retrieved from http://www.leeds.ac.uk/arts/homepage/387/leeds_working_papers_in_linguistics_and_phonetics

Koenig, L. L. (2000). Laryngeal factors in voiceless consonant production in men, women, and 5-year-olds. *Journal of Speech, Language, and Hearing Research, 43*, 1211–1228.

Kohler, K. (1977). *Einführung in die Phonetik des Deutschen*. Berlin: Schmidt.

Kupisch, T., & Rothman, J. (2016). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*. Advance online publication. doi: 10.1177/1367006916654355

Laeufer, C. (1996). The acquisition of a complex phonological contrast: Voice timing patterns of English initial stops by native French speakers. *Phonetica, 53*, 86–110.

Lisker, L. & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384–422.

Lleó, C., & Kehoe, M. (2002). On the interaction of phonological systems in child bilingual acquisition. *International Journal of Bilingualism, 6*, 233–237.

Macken, M. A., & Barton, D. (1980a). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language, 7*, 41–74.

Macken, M. A., & Barton, D. (1980b). The acquisition of the voicing contrast in Spanish: A phonetic and phonological study of word-initial stop consonants. *Journal of Child Language, 7*, 433–458.

MacLeod, A. A. (2017). Phonetic and phonological perspectives on the acquisition of voice onset time by French-speaking children. *Clinical Linguistics & Phonetics, 30*, 584–598.

Maddieson, I. (1997). Phonetic Universals. In J. Laver & W. J. Hardcastle (Eds.), *The handbook of phonetic sciences* (pp. 619–639). Oxford: Blackwell.

Major, R. C. (1987). English voiceless stop production by speakers of Brazilian Portuguese. *Journal of Phonetics, 15*, 197–202.

Mayr, R., Howells, G., & Lewis, R. (2015). Asymmetries in phonological development: The case of word-final cluster acquisition in Welsh-English bilingual children. *Journal of Child Language, 42*, 146–179.

Mayr, R., & Siddika, A. (2016). Inter-generational transmission in a minority language setting: Stop consonant production by Bangladeshi heritage children and adults. *International Journal of Bilingualism*. Advance online publication. doi: 10.1177/1367006916672590

McAllister Byun, T., Inkelas, S., & Rose, Y. (2016). The A-map model: Articulatory reliability in child-specific phonology. *Language, 92*, 141–178.

McCarthy, K. M., Mahon, M., Rosen, S., & Evans, B. G. (2014). Speech perception and production by sequential bilingual children: A longitudinal study of voice onset time acquisition. *Child Development, 85*, 1965–1980.

Menyuk, P., & Klatt, M. (1975). Voice onset time in consonant cluster production by children and adults. *Journal of Child Language, 2*, 223–231.

Mikuteit, S., & Reetz, H. (2007). Caught in the ACT: The timing of aspiration and voicing in East Bengali. *Language and Speech, 50*, 247–277.

Nearey, T. M., & Rochet, B. L. (1994). Effects of place of articulation and vowel context on VOT production and perception for French and English stops. *Journal of the International Phonetic Association, 24*, 1–18.

Neuhauser, S. (2011). Foreign accent imitation and variation of VOT and voicing in plosives. In W. S. Lee & E. Zee (Eds.), *Proceedings of the XVIIth International Congress of Phonetic Science* (pp. 1462–1465). Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/index.htm

Oller, D. K., & Eilers, R. E. (1982). Similarity of babbling in Spanish- and English-learning babies. *Journal of Child Language, 9*, 565–577.

Oller, D. K., Wieman, L. A., Doyle, W. J., & Ross, C. (1976). Infant babbling and speech. *Journal of Child Language, 3*, 1–11.

Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition*, *18*, 1–25.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America, 32,* 693–703.

Poon, P. G., & Mateer, C. A. (1985). A study of VOT in Nepali stop consonants. *Phonetica*, *42*, 39–47.

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

Riney, T., & Takagi, Y. (1999). Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning, 49*, 275–302.

Romaine, S. (1978). Postvocalic /r/ in Scottish English: Sound change in progress. In P. Trudgill (Ed.), *Sociolinguistic patterns in the British Isles* (pp. 144–157). London: Arnold.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, 25*, 421–436.

Schoonmaker-Gates, E. (2015). On voice-onset time as a cue to foreign accent in Spanish: Native and nonnative perceptions. *Hispania, 98*, 779–791.

Stock, D. (1971). *Untersuchungen zur Stimmhaftigkeit hochdeutscher Phonem-realisationen*. Hamburg: Buske.

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017). Second language attainment and first language attrition: The case of VOT in immersed Dutch-German late bilinguals. *Second Language Research, 33*, 483–518.

Szagun, G., Stumper, B., & Schramm, S. A. (2009). *FRAKIS - Fragebogen zur frühkindlichen Sprachentwicklung. PSYNDEX Tests Info*. Frankfurt am Main: Pearson.

Tamburelli, M., Sanoudaki, E., & Sowinska, M. (2015). Acceleration in the bilingual acquisition of phonological structure: Evidence from Polish-English bilingual children. *Bilingualism: Language and Cognition, 18*, 713–725.

Umeda, N. (1977). Consonant duration in American English. *Journal of the Acoustical Society of America, 61*, 846–858.

Unsworth, S. (2008). Age and input in the acquisition of grammatical gender in Dutch. *Second Language Research, 24*, 365–396.

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*, 86–110.

Unsworth, S., Argyri, E., Cornips, L., Hulk, A., Sorace, A., & Tsimpli, I. (2014). On the role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics, 35,* 765–805.

van Alphen, P. M. (2004). Perceptual relevance of prevoicing in Dutch. Ph.D. dissertation, Radboud University Nijmegen, The Netherlands.

van Alphen, P. M., & McQueen, J. M. (2006). The effect of voice onset time differences on lexical access in Dutch. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 178–196.

van Alphen, P. M., & Smits, R. (2004). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics, 32,* 455–491.

Volaitis, L. E., & Miller, J. L. (1992). Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *Journal of the Acoustical Society of America, 92*, 723–735.

Wells, J. C. (1982). *Accents of English.* Cambridge: University Press.

Whiteside, S. P., Dobbin, R., & Henry, L. (2003). Patterns of variability in voice onset time: A developmental study of motor speech skills in humans. *Neuroscience Letters, 347,* 29–32.

Williams, L. (1980). Phonetic variation as a function of second-language learning. In G. Yeni-Komshian, J. Kavanagh & C. Ferguson (Eds.), *Child Phonology: Volume 2. Perception* (pp. 185–215). New York, NY: Academic Press.

Yeni-Komshian, G., Caramazza, A., & Preston, M. S. (1977). A study of voicing in Lebanese Arabic. *Journal of Phonetics, 5*, 35–48.

Yu, V. Y., De Nil, L. F., & Pang, E. W. (2015). Effects of age, sex and syllable number on voice onset time: Evidence from children's aspirated stops. *Language and Speech, 58*, 152–167.

Zink, I., & Lejaegere, M. (2002). *N-CDIs: Lijsten voor communicatieve ontwikkeling. Aanpassing en hernormering van de MacArthur CDIs van Fenson et al. 1993*. Leuven: Acco.

Zlatin, M. A., & Koenigsknecht, R. A. (1976). Development of the voicing contrast: A comparison of voice onset time in stop perception and production. *Journal of Speech and Hearing Research, 19,* 93–111.

Appendix 1. Target words

Dutch target words

Word	Pronunciation	German translation	German pronunciation	English translation
bal	[ˈbɑɫ]	Ball	[ˈbal]	ball
bed	[ˈbɛt]	Bett	[ˈbɛt]	bed
beer	[ˈbeːr]	Bär	[ˈbɛːɐ̯]	bear
boom	[ˈboːm]	Baum	[ˈbaʊ̯m]	tree
boot	[ˈboːt]	Boot	[ˈboːt]	boat
buik	[ˈbœyk]	Bauch	[ˈbaʊx]	tummy
deur	[ˈdøːr]	Tür	[ˈtyːɐ̯]	door
dieren	[ˈdiːrə]	Tiere	[ˈtiːʀə]	animals
dokter	[ˈdɔktər]	Doktor	[ˈdɔktoːɐ̯]	doctor/physician
doos	[ˈdoːs]	Karton	[ˈkɐtɔŋ]	cardboard box
douche	[ˈduʃ]	Dusche	[ˈduːʃə]	shower
duim	[ˈdœym]	Daumen	[ˈdaʊmən]	thumb
kaas	[ˈkaːs]	Käse	[ˈkɛːzə]	cheese
kast	[ˈkɑst]	Schrank	[ˈʃʀaŋk]	cupboard
kikker	[ˈkɪkər]	Frosch	[ˈfʀɔʃ]	frog
kip	[ˈkɪp]	Huhn	[ˈhuːn]	chicken
koe	[ˈku]	Kuh	[ˈkuː]	cow
koning	[ˈkoːnɪŋ]	König	[ˈkøːnɪç]	king
paard	[ˈpaːrt]	Pferd	[ˈp͡feːɐ̯t]	horse
pan	[ˈpɑn]	Topf	[ˈtɔp͡f]	pan
peer	[ˈpeːr]	Birne	[ˈbɪɐ̯nə]	pear
pink	[ˈpɪŋk]	kleiner Finger	[ˈklaɪ̯nɐˈfɪŋɐ]	pinky finger
pizza	[ˈpidza]	Pizza	[ˈpɪt͡sa]	pizza
pop	[ˈpɔp]	Puppe	[ˈpʊpə]	doll
taart	[ˈtaːrt]	Torte	[ˈtɔɐ̯tə]	pie
tafel	[ˈtaːfəɫ]	Tisch	[ˈtɪʃ]	table
tak	[ˈtɑk]	Zweig	[ˈt͡svaɪ̯k]	branch
tas	[ˈtɑs]	Tasche	[ˈtaʃə]	bag
tent	[ˈtɛnt]	Zelt	[ˈt͡sɛlt]	tent
tijger	[ˈtɛiɣər]	Tiger	[ˈtiːgɐ]	tiger

2

German target words

Word	Pronunciation	Dutch Translation	Dutch pronunciation	English translation
Ball	[ˈbal]	bal	[ˈbɑɫ]	ball
Bär	[ˈbɛːɐ̯]	beer	[ˈbeːr]	bear
Baum	[ˈbaʊm]	boom	[ˈboːm]	tree
Bett	[ˈbɛt]	bed	[ˈbɛt]	bed
Biene	[ˈbiːnə]	bij	[ˈbɛi]	bee
Birne	[ˈbɪɐ̯nə]	peer	[ˈpeːr]	pear
Dach	[ˈdax]	dak	[ˈdɑk]	roof
Daumen	[ˈdaʊmən]	duim	[ˈdœym]	thumb
Decke	[ˈdɛkə]	deken	[ˈdekən]	blanket
Doktor	[ˈdɔktoːɐ̯]	dokter	[ˈdɔktər]	doctor/physician
Dose	[ˈdoːzə]	potje	[ˈpɔtjə]	box
Dusche	[ˈduːʃə]	douche	[ˈduʃ]	shower
Käse	[ˈkɛːzə]	kaas	[ˈkaːs]	cheese
Katze	[ˈkat͡sə]	kat / poes	[ˈkɑt] / [ˈpuːs]	cat
Kette	[ˈkɛtə]	ketting	[ˈkɛtɪŋ]	necklace
Korb	[ˈkɔɐ̯p]	mand	[ˈmɑnt]	basket
Kuh	[ˈkuː]	koe	[ˈku]	cow
Küken	[ˈkyːkən]	kuiken	[ˈkœykən]	chick
Pilz	[ˈpɪl͡ts]	paddenstoel	[ˈpɑdənstuːɫ]	mushroom
Pinsel	[ˈpɪnzəl]	kwast	[ˈkwɑst]	paint brush
Pizza	[ˈpɪt͡sa]	pizza	[ˈpidza]	pizza
Pommes	[ˈpɔməs]	frites	[ˈfrit]	French fries
Puppe	[ˈpʊpə]	pop	[ˈpɔp]	doll
Puzzle	[ˈpʊzəl]	puzzel	[ˈpʏzəɫ]	jigsaw
Tasse	[ˈtasə]	kop	[ˈkɔp]	cup
Teller	[ˈtɛlɐ]	bord	[ˈbɔrt]	plate
Tiere	[ˈtiːʀə]	dieren	[ˈdiːrə]	animals
Tiger	[ˈtiːgɐ]	tijger	[ˈtɛiɣər]	tiger
Tisch	[ˈtɪʃ]	tafel	[ˈtaːfəɫ]	table
Tür	[ˈtyːɐ̯]	deur	[ˈdøːr]	door

Appendix 2. Analysis by vocalic context

Voiceless plosives: mean VOT values (ms) by place of articulation and vocalic context by language and language background over children.

	Dutch			German		
	Monolingual	Bilingual	Number of target words	Monolingual	Bilingual	Number of target words
/p/						
open vowel	13	23	2	–	–	0
open-mid vowel	21	22	1	61	36	1
close-mid vowel	16	24	1	–	–	0
close vowel	14	17	2	51	31	5
/t/						
open vowel	26	22	4	60	29	1
open-mid vowel	25	34	2	67	44	1
close-mid vowel	–	–	0	–	–	0
close vowel	–	–	0	84	66	4
/k/						
open vowel	24	33	2	66	47	1
open-mid vowel	–	–	0	74	53	3
close-mid vowel	35	39	1	–	–	0
close vowel	38	44	3	82	68	2

Due to constraints in the selection of target words, no even distribution of open and close vowels across consonantal places of articulation and languages could be achieved. This imbalance does not affect the two analyses comparing the bilinguals to monolinguals in Dutch and in German, as all participants produced the same target words.

The uneven distribution of vocalic contexts could potentially be a conflict in the comparison of the bilinguals' VOT across Dutch and German. To address this concern, we ran an analysis on /k/, which is the only consonantal place of articulation for which the distribution of open and close vowels is approximately even in Dutch and German. For /k/, we have two open and four close (including close-mid) vowels in Dutch and four open (including open-mid) and two close vowels in German. If the language differentiation we observed in the bilinguals were caused by the different distribution of the vocalic contexts instead of being a language effect, we would expect longer VOT in Dutch than in German. However, the analysis shows that despite the larger amount of target words with close vowels in Dutch, the bilingual children produced longer VOT in German /k/ (M=57 ms) than in Dutch /k/ (M=40 ms; β=10.65, SE=3.97, t=2.68, p=.007). Based on these results, we are confident that our finding – that bilinguals produce longer VOT in German than in Dutch – indeed indicates language differentiation and does not result from differences in vocalic contexts between the stimuli used for the two languages.

Appendix 3. Output of the statistical models

VOT productions of Dutch and German monolingual children

Voiceless plosives (2308 observations)				
	β	*SE*	*t*	*p*
Intercept	51.953	2.950	17.610	<.001
Language	28.962	2.950	9.817	<.001
Age	0.174	0.145	1.203	.230
Task	-1.402	0.680	-2.063	.039
PoA (labial - coronal)	-13.942	3.461	-4.028	<.001
PoA (coronal - dorsal)	3.296	3.572	0.923	>.250
Word Length	4.493	1.364	3.293	<.001
Language*Age	0.101	0.145	0.694	>.250
Language*Task	1.269	0.680	1.866	.062
Language*PoA (labial - coronal)	-7.283	3.461	-2.104	.036
Language*PoA (coronal - dorsal)	-6.210	3.572	-1.738	.082
Language*Word Length	4.334	1.364	3.177	.002
Voiced plosives (1475 observations)				
	β	*SE*	*t*	*p*
Intercept	1.657	0.194	8.543	<.001
Language	1.564	0.194	8.068	<.001
Age	-0.022	0.019	-1.141	>.250
Task	-0.287	0.109	-2.632	.009
PoA	-0.546	0.105	-5.215	<.001
Language*Age	0.002	0.019	0.108	>.250
Language*Task	0.155	0.109	1.421	.156
Language*PoA	-0.212	0.105	-2.028	.042

VOT productions of bilingual children in Dutch and in German (Research Question 1)

Voiceless plosives (2378 observations)				
	β	*SE*	*t*	*p*
Intercept	44.414	4.667	9.517	<.001
Language	14.426	3.537	4.078	<.001
Age	-0.227	0.285	-0.795	>.250
Exposure	0.346	0.158	2.186	.029
Task	-2.209	0.957	-2.309	.021
PoA (labial - coronal)	-16.817	4.569	-3.681	<.001
PoA (coronal - dorsal)	4.688	4.658	1.006	>.250
Word Length	0.584	1.673	0.349	>.250
Language*Age	-0.181	0.162	-1.121	>.250
Language*Exposure	0.256	0.090	2.851	.004
Language*Task	1.074	0.485	2.213	.027
Language*PoA (labial - coronal)	-7.986	4.442	-1.798	.072
Language*PoA (coronal - dorsal)	-4.833	4.411	-1.096	>.250
Language*Word Length	2.226	1.639	1.358	.174
Voiced plosives (1467 observations)				
	β	*SE*	*t*	*p*
Intercept	1.473	0.305	4.833	<.001
Language	0.260	0.173	1.497	.134
Age	-0.022	0.031	-0.721	>.250
Exposure	0.022	0.018	1.224	.221
Task	-0.375	0.122	-3.063	.002
PoA	-0.322	0.115	-2.812	.005
Language*Age	0.016	0.016	1.009	.313
Language*Exposure	0.022	0.021	1.22	.223
Language*Task	-0.151	0.078	-1.926	.054
Language*PoA	-0.140	0.102	-1.379	.168

VOT productions of bilingual and monolingual children in Dutch (Research Question 2)

Voiceless plosives (2224 observations)				
	β	*SE*	*t*	*p*
Intercept	25.915	3.099	8.362	<.001
Language Background	2.862	1.915	1.495	.134
Age	0.089	0.143	0.624	>.250
Task	-2.960	0.609	-4.858	<.001
PoA (labial - coronal)	-5.942	3.645	-1.630	.103
PoA (coronal - dorsal)	10.369	3.672	2.824	.005
Word Length	-1.047	1.326	-0.790	>.250
Language Background*Age	0.151	0.143	1.061	>.250
Language Background*Task	-0.208	0.516	-0.403	>.250
Language Background*PoA(labial - coronal)	-0.394	1.586	-0.249	>.250
Language Background*PoA(coronal - dorsal)	0.252	1.723	0.147	>.250
Language Background*Word Length	-0.022	0.545	-0.041	>.250
Voiced plosives (1350 observations)				
	β	*SE*	*t*	*p*
Intercept	0.640	0.214	2.988	.003
Language Background	0.510	0.209	2.403	.016
Age	-0.034	0.022	-1.574	.116
Task	-0.352	0.105	-3.368	<.001
PoA	-0.234	0.102	-2.290	.022
Language Background*Age	-0.011	0.022	-0.502	>.250
Language Background*Task	0.110	0.102	1.079	>.250
Language Background*PoA	0.077	0.090	0.846	>.250

VOT productions of bilingual and monolingual children in German (Research Question 2)

Voiceless plosives (2462 observations)				
	β	*SE*	*t*	*p*
Intercept	69.619	5.275	13.198	<.001
Language Background	-10.200	3.121	-3.268	.001
Age	-0.146	0.252	-0.579	>.250
Task	-0.616	0.855	-0.721	>.250
PoA (labial - coronal)	-23.445	6.353	-3.690	<.001
PoA (coronal - dorsal)	-0.975	6.399	-0.152	>250
Word Length	5.857	1.945	3.012	.003
Language Background*Age	-0.164	0.252	-0.649	>.250
Language Background*Task	-0.498	0.853	-0.583	>.250
Language Background*PoA(labial - coronal)	-0.669	2.176	-0.307	>.250
Language Background*PoA(coronal - dorsal)	0.996	2.365	0.421	>.250
Language Background*Word Length	-0.876	0.918	-0.954	>.250
Voiced plosives (1592 observations)				
	β	*SE*	*t*	*p*
Intercept	2.584	0.245	10.544	<.001
Language Background	-0.944	0.254	-3.720	<.001
Age	-0.025	0.024	-1.047	>.250
Task	-0.212	0.114	-1.858	.063
PoA	-0.728	0.118	-6.166	<.001
Language Background*Age	-0.005	0.024	-0.187	>.250
Language Background*Task	-0.199	0.110	-1.806	.071
Language Background*PoA	0.269	0.134	2.012	.044

Appendix 4. Post-hoc analyses on interactions.

Monolingual Dutch children and monolingual German children

The main model comparing the monolingual Dutch and German children's VOT of 'voiceless' plosives revealed a Language x Place of Articulation (/p/ vs. /t/) interaction and a Language x Word Length (monosyllabic vs. disyllabic words) interaction. Several post-hoc analyses followed up on these interactions. The first set of analyses confirmed that the effect of Language is significant in both /p/ (β=18.28, *SE*=2.62, *t*=6.99, *p*<.001) and /t/ (β=27.39, *SE*=2.96, *t*=9.24, *p*<.001) and also in both monosyllabic words (β=32.83, *SE*=3.90, *t*=8.41, *p*<.001) and disyllabic words (β=24.35, *SE*=4.09, *t*=5.95, *p*<.001).

Subsequent analyses, conducted separately by language, showed that only the German monolingual children produced shorter VOT in /p/ than in /t/ (German: β=-21.34, *SE*=5.89, *t*=-3.36, *p*<.001; Dutch: β=-6.33, *SE*=3.60, *t*=-1.76, *p*=.078) and shorter VOT in disyllabic than in monosyllabic words (German: β=8.79, *SE*=2.37, *t*=3.71, *p*<.001; Dutch: β=0.085, *SE*=1.31, *t*=0.06, *p*>.250).

The analysis comparing the monolingual children's percentage of prevoiced 'voiced' plosives similarly revealed a Language x Place of Articulation interaction. Post-hoc analyses confirmed that the effect of Language is significant in /b/ (β=1.36, *SE*=0.18, *z*=7.46, *p*<.001) and also in /d/ (β=1.87, *SE*=0.28, *z*=6.63, *p*<.001).

Subsequent by-language analyses indicated that both the Dutch monolingual children and the German monolingual children prevoiced /b/ more frequently than /d/, but the magnitude of the effect was larger in the German monolingual children (β=-1.06, *SE*=0.23, *z*=-4.52, *p*<.001) than in the Dutch monolingual children (β=-0.32, *SE*=0.12, *z*=-2.80, *p*=.005).

Bilingual children in Dutch and in German (Research question 1)

The main model comparing the VOT of 'voiceless' plosives in the bilingual children's Dutch and German revealed a Language x Exposure to German interaction and a Language x Elicitation Task interaction. The first was explored in by-language post-hoc analyses and revealed that children with more exposure to German produced longer VOT in German (β=0.52, *SE*=0.24, *t*=2.17, *p*=.030), but not in Dutch (β=0.15, *SE*=0.14, *t*=1.09, *p*>.250). For the Language x Elicitation Task interaction, post-hoc analyses first confirmed that the effect of Language is significant both in the story elicitation task (β=13.54, *SE*=3.59, *t*=3.77, *p*<.001) and in the game elicitation task (β=16.27, *SE*=3.71, *t*=4.39, *p*<.001). Separate by-language analyses showed that the bilingual children only produced longer VOT in the story task than in the game task when they spoke Dutch (β=-3.14, *SE*=0.92, *t*=-3.42, *p*<.001), but not when they spoke German (β=-1.11, *SE*=1.25, *t*=-0.89, *p*>.250).

Bilingual children and monolingual German children (Research question 2)

The main model comparing the bilinguals' and monolinguals' percentage of prevoiced 'voiced' plosives in German revealed a Language Background x Place of Articulation interaction. Post-hoc analyses confirmed that the effect of Language Background is significant in both /b/ (β=-0.73, *SE*=0.24, *z*=-3.08, *p*=.002) and /d/ (β=-1.32, *SE*=0.37, *z*=-3.54, *p*<.001). Separate by-language analyses indicated that the bilingual children as well as the monolingual children prevoiced /b/ more frequently than /d/, but the magnitude of the effect was larger in the monolingual children (β=-1.06, *SE*=0.23, *z*=-4.52, *p*<.001) than in the bilingual children (β=-0.41, *SE*=0.15, *z*=-2.79, *p*=.005).

Chapter 3

Bilingual preschoolers' voicing perception supports language-specific voicing systems

Based on:

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. *Bilingual preschoolers' voicing perception supports language-specific voicing systems.* Manuscript in preparation.

3

Abstract

Bilingual children's *production* of the voicing contrast has received substantial attention, and research has revealed that bilingual children generally produce 'voiceless' plosives with distinct voice onset time (VOT) between their two languages. By contrast, bilingual children seem to produce language-general VOT for 'voiced' plosives, which leaves the conclusion of fully language-specific voicing systems contested. In the present study, we investigated whether Dutch-German simultaneous bilingual children aged 3;9 to 5;11 years *perceive* VOT language-specifically. An XAB speech sound categorization task with acoustically synthesized CV-syllables was conducted in a Dutch session and in a German session. The results show that the bilingual children associate 'voiced' plosives with longer VOT values in German than in Dutch, which supports that voicing systems are language-specific in simultaneous bilingual acquisition.

1 Introduction

The nature of bilinguals' voicing representations has been a central topic in bilingualism research since the 1970s, and this research provided evidence that early bilingual adults differentiate the phonetic realization of voicing between their two languages in both production (Lein, Kupisch & van de Weijer, 2016; MacLeod & Stoel-Gammon, 2009, 2010; Sundara, Polka & Baum, 2006) and perception (Elman, Diehl & Buchwald, 1977; Flege & Eefting, 1987; Garcia-Sierra, Diehl & Champlin, 2009; Gonzales & Lotto, 2013; Hazan & Boulakia, 1993). Whether these language-specific voicing systems are already in place in early childhood remains controversial.

When bilingual children acquire two languages that differ in their phonetic implementation of the voicing contrast, they are confronted with the challenge to acquire language-specific voicing categories in each of their languages. Previous research found that bilingual children produce 'voiceless' plosives differently between their two languages, while their productions of 'voiced' plosives do not typically differ between their languages. These seemingly contradictory findings leave the conclusion of fully language-specific voicing systems contested (Deuchar & Clark, 1996; Johnson & Wilson, 2002 (the older child); Khattab, 2000; Mayr & Siddika, 2016). The targeted languages in these studies were English (the language of the environment) in addition to either Spanish, Arabic, Japanese or Sylheti (the heritage language). Children across studies resorted to an English-like realization of 'voiced' plosives in both their languages.

The discrepancies between the acquisition outcomes for 'voiced' and 'voiceless' plosives in bilingual first language acquisition may be related to acoustic saliency: the difference between 'voiced' plosives of the investigated languages is acoustically less salient than the difference between the languages' 'voiceless' plosives (García-Sierra, Ramírez-Esparza, Silva-Pereyra, Siard & Champlin, 2012). Due to a lack of an acoustically salient difference between the two languages' 'voiced' plosives, bilingual children may not acquire language-specific representations for these categories. The finding that bilingual children generally resort to an English-like realization of 'voiced' plosives may be related to the overall stronger influence of the language of the environment compared to the heritage language.

An alternative explanation for bilingual children's productions of English-like 'voiced' plosives in both languages is related to developmental constraints. In English, 'voiced' plosives typically have short lag voice onset time (VOT), which means that burst release and onset of vocal fold vibration approximately coincide.

In languages such as Spanish, Arabic, Japanese, and Sylheti, 'voiced' plosives are prevoiced, which requires initiation and sustainment of vocal fold vibration prior to burst release (Kewley-Port & Preston, 1974). The realization of short lag VOT is aerodynamically easier than the production of prevoicing and these differences are reflected in monolingual children's acquisition of the voicing contrast. Monolingual children acquiring English produce target-like 'voiced' plosives from early on, while children acquiring French, Spanish, and Arabic only consistently produce target-like 'voiced' plosives in the early school years (Kewley-Port & Preston, 1974; Khattab, 2000; Macken & Barton, 1980a, 1980b; MacLeod, 2016). If children's productions of 'voiced' plosives deviate from the adult target because of developmental constraints, no firm conclusions can be drawn about their voicing representations.

The language-specific production pattern for 'voiceless' plosives in both languages in combination with seemingly similar productions of 'voiced' plosives has recently been confirmed in bilingual children who acquired German as a heritage language in the Netherlands (Stoehr, Benders, van Hell & Fikkert, 2017b / Chapter 2). Interestingly, these children produced 'voiced' plosives with the more complex Dutch-like prevoicing, which rules out that bilingual children automatically opt for the developmentally simpler realization. Yet, the consistent finding that no language differentiation is observed in bilingual children's production of 'voiced' plosives could mean that bilingual children do not have language-specific categories for 'voiced' plosives; but it could also be the case that there is a language-specific difference between the 'voiced' categories in the two languages, which remained undetected.

Investigations of bilingual children's voicing *perception* can answer whether bilingual children already have language-specific voicing systems similar to that of early bilingual adults. If bilingual children perceive voicing specific to each of their native languages, this would provide evidence in favor of language-specific voicing systems even though their voicing productions in the two languages overlap. The present study is designed to answer this open question using the same participants whose voicing productions have recently been reported (Stoehr et al., 2017b / Chapter 2).

1.1 Voicing perception

The phonological voicing contrast is composed of several acoustic cues. For word- and syllable-initial plosives preceding stressed vowels, these include voice onset time (VOT), aspiration amplitude, relative amplitude of the release burst, fundamental frequency (F0) contour, first formant (F1) onset and transition, and the duration of the following vowel (Forrest & Rockman, 1988; Haggard, Ambler & Callow, 1970; Liberman, Delattre & Cooper, 1957; Lisker, 1975; Lisker & Abramson, 1964; Repp, 1979; Summerfield, 1981; Summerfield & Haggard, 1977).

VOT has been identified as the most important cue to voicing and listeners can make voicing judgements based on VOT alone (Abramson & Lisker, 1973; Lisker, 1978; Lisker & Abramson, 1970). VOT is the time between the release of a plosive, such as opening of the lips in /p/, and the onset of vocal fold vibration. Across languages, the exact VOT values used to contrast voicing can differ. Languages such as Dutch, Arabic, French, Japanese, Spanish, and Sylheti, have 'voiced' plosives in which vocal fold vibration starts prior to consonantal release, resulting in negative VOT values referred to as *prevoicing*. The 'voiceless' counterpart category in these languages is realized with short positive VOT, henceforth *short lag* VOT, resulting from an onset of vocal fold vibration concurrent with or shortly after consonantal release. By contrast, German and English 'voiced' plosives are produced with similar short lag VOT as 'voiceless' plosives in Dutch and other prevoicing languages. 'Voiceless' plosives in German and English have a substantial delay in the onset of vocal fold vibration relative to consonantal release, resulting in long positive VOT values called *aspiration*. Bilinguals who acquire language combinations such as Dutch and German have to resolve the phonological overlap in the short lag VOT range in their perception and production in order to perform on par with their monolingual peers.

Speech sound perception tasks allow establishing cross-linguistic differences in the perception of VOT. These tasks are based on the concept that listeners perceive an acoustic continuum as distinct sound categories, a phenomenon called categorical perception. The perception of variation of an acoustic parameter along a continuum rapidly changes from one discrete category to the other once the category boundary has been crossed (Liberman, Harris, Hoffman & Griffith, 1957). With regards to VOT, monolingual English listeners' perception of voicing rapidly changes from 'voiced' to 'voiceless' between 20 ms and 30 ms of VOT (Lisker & Abramson, 1970).

The first studies testing adult bilinguals' voicing perception found similar category boundary locations in both languages, and thus did not provide evidence for language-specific voicing systems, possibly due to a lack of language context in the experimental paradigm (Caramazza, Yeni-Komshian, Zurif & Carbone, 1973; Williams, 1977). By contrast, experiments which set a clear language context provided evidence for language-specific voicing systems in early adult bilinguals by showing that French-English and Spanish-English bilinguals associate 'voiced' plosives with longer VOT values in English than in French or Spanish (Elman et al., 1977; Garcia-Sierra et al., 2009; Gonzales & Lotto, 2013; Hazan & Boulakia, 1993). The language context in these studies was either set by using real words in both languages (Hazan & Boulakia, 1993), pseudowords with language-specific phonemes in the second syllable (Gonzales & Lotto, 2013), or CV syllables regularly accompanied by auditory presentation of sentences in the test language (Elman et al., 1977; Garcia-Sierra et al., 2009).

The first evidence for early emergence of language-specific voicing systems comes from bilingually raised preverbal infants' brain responses to CV syllables manipulated in VOT (Ferjan Ramírez, Ramírez, Clarke, Taulu & Kuhl, 2016; Garcia-Sierra, Rivera-Gaxiola, Percaccio, Conboy, Romo, Klarman, Ortiz & Kuhl, 2011). In an EEG/ERP study using the mismatch negativity (MMN) paradigm, Spanish-English bilingually raised infants aged 10-12 months showed different peak amplitudes to standard versus deviant stimuli manipulated in VOT in both English and Spanish (Garcia-Sierra et al., 2011)[7]. In a whole head magnetoencephalography study, 11-months old Spanish-English bilingually raised infants and monolingual English infants showed a similar mismatch response (MMR) to the English voicing contrast, but only the bilingual infants also showed an MMR to the Spanish voicing contrast (Ferjan Ramírez et al., 2016). These studies suggest that by the end of the first year of life, infants exposed to English and Spanish show neural sensitivity to the voicing contrasts of both languages. Assuming that children build phonological representations in their mental lexicon in the course of word learning, it is crucial to follow up on the results of Ferjan Ramírez et al. (2016) and Garcia-Sierra et al. (2011) with bilingual children past the vocabulary spurt (Fikkert, 2010; Werker & Curtin, 2005).

To date, no data are available on older early bilingual children's voicing perception in both of their languages after they started producing words, but one study has tested preschool-aged early bilingual children's voicing perception in their majority language (McCarthy, Mahon, Rosen and Evans, 2014). McCarthy et al. tested the perception of the English voicing contrast in Sylheti-English bilingual children at two time points, once when they were aged 4;4, and again at 5;4. At the first time of testing, the bilinguals had only been exposed to English for seven months. The location of the voicing boundary between bilinguals and monolinguals

[7] Adopting an α-level of .05, the differences in peak amplitude between standard and deviant in Spanish were not statistically significant (*p*=.06), and were interpreted as "marginally significant" by Garcia-Sierra et al. (2011).

3

did not differ detectably, but the bilinguals had a shallower categorization slope than monolinguals, which means that the bilinguals categorized the VOT stimuli close to the category boundary less consistently. One year later, the bilingual and monolingual children's categorization slopes did not differ anymore in steepness. As McCarthy et al. (2014) did not test the children's voicing perception in Sylheti, their study does not provide further insight into the language-specificity of bilinguals' voicing systems. The study showed, however, that bilingual children's voicing system in the language of the environment develops toward more monolingual-like perception with increasing language exposure and age.

In monolingual first language acquisition research, speech sound categorization tasks have frequently been used to investigate the perception of syllable-initial voicing in children as young as two years of age (Bailey & Haggard, 1980; Burnham, Earnshaw & Clark, 1991; Giezen, 2011; Hazan & Barrett, 2000; Hoonhorst, Medina, Colin, Markessis, Radeau, Deltenre & Serniclaes, 2011; McCarthy et al, 2014; Simon & Fourcin, 1978; Wolf, 1973; Zlatin & Koenigsknecht, 1975). These studies collectively suggest that children rely on VOT as a cue to voicing in Dutch, English, and French. Between the ages of two and six years, children's category boundaries become steeper and therefore more adult-like, pointing to increased consistency in mapping VOT onto phonological categories. This category boundary sharpening seems to continue even in the second decade of life (Hazan & Barrett, 2000).

Cross-sectional and cross-linguistic research on English and French children aged 2 to 14 years has identified three developmental stages in children's labeling behavior in speech sound categorization tasks, in which category boundaries become steeper as a function of age (Simon & Fourcin, 1977). *Scattered labeling* describes confident labeling of stimuli which match the actual linguistic environment in terms of VOT, while stimuli with intermediate VOT values appear to be labeled randomly. Scattered labeling had been observed in most English-speaking children aged two to three years and French-speaking children aged two to four years. *Progressive labeling* defines a trend in which responses from one category to the other increase linearly. This pattern occurred mostly between three to four years of age in English-speaking children and persisted up to seven years of age in French-speaking children. Adult-like *categorical labeling,* which is characterized as a rapid change from one category to the other, was present after the age of five to six years in English-speaking children, and started about one to two years later for French-speaking children.

In sum, speech sound categorization tasks are adequate and frequently used measures of voicing perception in monolingual children as well as monolingual and bilingual adults. Testing children in a speech sound categorization task requires a child-friendly set-up of the task (Bailey & Haggard, 1980; Brasileiro, 2009; Burnham et al., 1991; Giezen, 2011; Hoonhorst et al., 2011; McCarthy et al., 2014; Simon & Fourcin, 1978; Wolf, 1973; Zlatin & Koenigsknecht, 1975). Moreover, when testing bilinguals, it is essential to set a clear language context during the experimental task (Elman et al., 1977; Garcia-Sierra et al., 2009; Gonzales & Lotto, 2013; Hazan & Boulakia, 1993).

1.2 The present study

The present study is designed to test the perception of voicing in both languages of simultaneous bilingual preschoolers speaking Dutch and German and their monolingual peers. To focus on the role of VOT in voicing perception, stimulus continua ranging from [ba] to [pʰa], [da] to [tʰa], and [ga] to [kʰa] were acoustically synthesized. Experiment I was specifically designed to obtain Dutch and German adult native speakers' category boundaries for the stimuli used in this study. Subsequently, Experiment II turns to the perception of voicing in bilingual children.

The bilingual children grew up in the Netherlands where they acquired Dutch as the majority language and German as the heritage language from one or both parents. The amount of exposure to each language has previously been identified as an important factor in bilingual children's VOT production (Stoehr et al., 2017b / Chapter 2) and in bilingual infants' brain responses to voicing contrasts (Garcia-Sierra et al., 2011), and is therefore used as an independent variable in the present research. Based on these previous studies, it is expected that bilingual children with more exposure to the heritage language German are more likely to perceive voicing differently in German than in Dutch. The central questions are whether Dutch-German simultaneous bilingual preschoolers differ in the location and steepness of their voicing boundary in Dutch and in German, and whether language-specific voicing boundary as well as boundary steepness are associated with current heritage language exposure. We furthermore ask whether bilingual children differ from their monolingual peers in their voicing perception.

If bilingual preschoolers have language-specific voicing systems, their perceptual boundary between 'voiced' and 'voiceless' plosives is expected to be located on lower VOT values in Dutch than in German. With respect to language exposure, it is possible that only children with balanced bilingual input have language-specific category boundaries for voicing or that children with little exposure to one language have a shallower category boundary in that language. If bilingual preschoolers do not yet have language-specific voicing systems, their category boundary is expected to be either intermediate to those of the respective monolingual children, or perhaps more aligned with the Dutch boundary location, given that the children tested in this study grow up in the Netherlands.

2 Experiment I: Adults

2.1 The PAM model and research questions

The aim of the present experiment is two-fold. The first aim is to establish that the VOT stimuli that were newly synthesized for the present study can be categorically perceived for voicing by Dutch and German adult listeners, and that the voicing boundary locations of Dutch and German adult native listeners differ. These two factors are important prerequisites for Experiment II. Moreover, German contrasts voicing at the labial, coronal and dorsal place of articulation, while Dutch lacks the ‘voiced’ dorsal plosive /g/. The second aim of the present study is to test whether Dutch listeners generalize the voicing contrast to the dorsal place of articulation.

The perception of non-native speech sounds, such as the ‘voiced’ dorsal plosive /g/ in Dutch, can be captured by the Perceptual Assimilation Model (PAM; Best, 1995). The PAM assumes that non-native segments are processed according to their similarities to native segments, resulting in two primary perception patterns: 1) adoption as a novel segment that may share features with native segments, or 2) assimilation to the perceptually closest native segment. The third pattern within PAM is perception as a non-speech sound, but given that /g/ shares features with other Dutch phonemes this is an unlikely prediction for the present experiment.

If Dutch native speakers perceive /g/ as a novel segment, they are expected to have a category boundary between /g/ and /k/ that is similar in steepness to their boundary at the labial and coronal places of articulation. If they map /g/ to the perceptually closest native category, they are expected to perceive /g/ as /k/, with which it shares manner and place of articulation, or /d/ or /b/, with which it shares manner of articulation and voicing.

2.2 Method

2.2.1 Participants

Thirty-two adult native speakers of Dutch (N=16, M_{age}=23 years, SD_{age}=4.70 years, range=18-35 years, 14 female) and German (N=16, M_{age}=22.25 years, SD_{age}=4.46 years, range=15-32 years, 13 female) participated in the experiment. One additional native speaker of German had been tested, but was excluded from the analysis because of low accuracy during the practice phase, as explained in section 2.2.3. Dutch speakers were tested in Nijmegen in the Central Eastern Netherlands and German speakers were tested in Cologne in Central Western Germany. All participants reported passive L2 proficiency, but all of them indicated that their daily language use was limited to their native language. They further reported being most comfortable using their native language, and rated their native language as their dominant language.

2.2.2 Materials

Stimuli were acoustically synthesized CV syllables composed of either a labial, coronal or dorsal plosive with manipulated VOT ranging from -80 to +80 ms, and the vowel /a/ as nucleus. The remainder of this section provides detailed information on the synthesis process.

Acoustic synthesis was done in the KlattWorks interface (McMurray, in preparation) of the Klatt (1980) synthesizer using a sampling rate of 11,025 Hz. The user interface of KlattWorks allows managing acoustic parameters, and provides parameterized linear and logistic functions for specifying individual parameters over time (McMurray, in preparation).

Labial, coronal and dorsal plosives were acoustically synthesized by a manipulation of the formant transitions and the quality and duration of the burst, modeled on tokens of natural speech. A set of 33 stimuli per place of articulation were manipulated in 5 ms VOT steps ranging from 80 ms before the burst to 80 ms after the burst.

During the burst of unaspirated plosives (-80 ms to 0 ms), the parameter Amplitude of Voicing (AV) was set to 25 dB, whereas the burst of aspirated plosives (+5 ms to 80 ms) was unvoiced. All other parameters were kept constant. Prevoicing was synthesized using the parameters Amplitude of Parallel Branch Voicing (AVS) and Amplitude of Nasality (AN), which were set to 50 dB. The duration of prevoicing was manipulated by cutting back these parameters from burst onset (i.e., 0 ms) to -80 ms. No other parameters were used during this timeframe. Aspiration was synthesized with the parameter Amplitude of Aspiration (AH), which was set to 30 dB from 5 ms after the burst offset until the end of the duration of aspiration. The duration of aspiration was manipulated by changing the end time of AH, while keeping all other parameters constant.

Formant trajectories and burst durations differed by consonantal place of articulation. Formant trajectories were rising for labials (F1-F3); rising (F1) and falling (F2-F3) for coronals; and rising (F1 and F3) and falling (F2) for dorsals (Figure 1). Transitions had a slope of +100 or -100, which resulted in transition durations of approximately 50 ms. Each transition had its midpoint 10 ms after the offset of the burst. The fourth to sixth formant had no transitions. The duration of the burst was 5 ms for labial stimuli, and 15 ms for coronals and dorsals. Table 1 displays all burst parameters by place of articulation.

The duration of the vowel [a] was set to 300 ms. The formant values of the vowel were set as follows throughout the vowel, F1= 900 Hz, F2= 1500 Hz, F3= 2600 Hz, F4= 4100 Hz, F5= 4558 Hz, and F6= 4990 Hz. F0 started at 260 Hz and gradually fell to 180 Hz. Intensity fell from 55 dB to 45 dB throughout the vowel.

Table 1. Burst parameters by place of articulation.

Parameter	Labial	Coronal	Dorsal
Amplitude of Frication (AF)	70 dB	55 dB	60 dB
Amplitude of Frication-Excited Parallel Bypass (AB)	52 dB	50 dB	60 dB
Amplitude of Parallel Tract 1st Formant (A1)	10 dB	30 dB	5 dB
Amplitude of Parallel Tract 2nd Formant (A2)	10 dB	30 dB	55 dB
Amplitude of Parallel Tract 3rd Formant (A3)	10 dB	30 dB	55 dB
Amplitude of Parallel Tract 4th Formant (A4)	30 dB	40 dB	55 dB
Amplitude of Parallel Tract 5th Formant (A5)	30 dB	45 dB	15 dB
Amplitude of Parallel Tract 6th Formant (A6)	20 dB	35 dB	5 dB
Duration of the burst	5 ms	15 ms	15 ms
Initial F1	350 Hz	450 Hz	300 Hz
Initial F2	850 Hz	1800 Hz	1750 Hz
Initial F3	2200 Hz	3000 Hz	2000 Hz

2.2.3 Design and procedure

Prior to testing, all participants gave informed consent[8] and completed a language history questionnaire. A multiple forced choice task programmed in Praat (Boersma & Weenink, 2014) was administered on a Macintosh notebook. Auditory stimuli were presented over Sennheiser HD 280 64 ohm headphones. Orthographic representations of the six response categories were presented on the computer screen. Half of the participants saw the response categories <ba>, <da>, and <ga>, corresponding to the 'voiced' plosives, on the left side of the screen, and the other half saw these response categories on the right side of the screen. Participants indicated their responses by clicking on the recognized response category with a

[8] One minor participated in the study, and consent was given by the participant and the legal guardians.

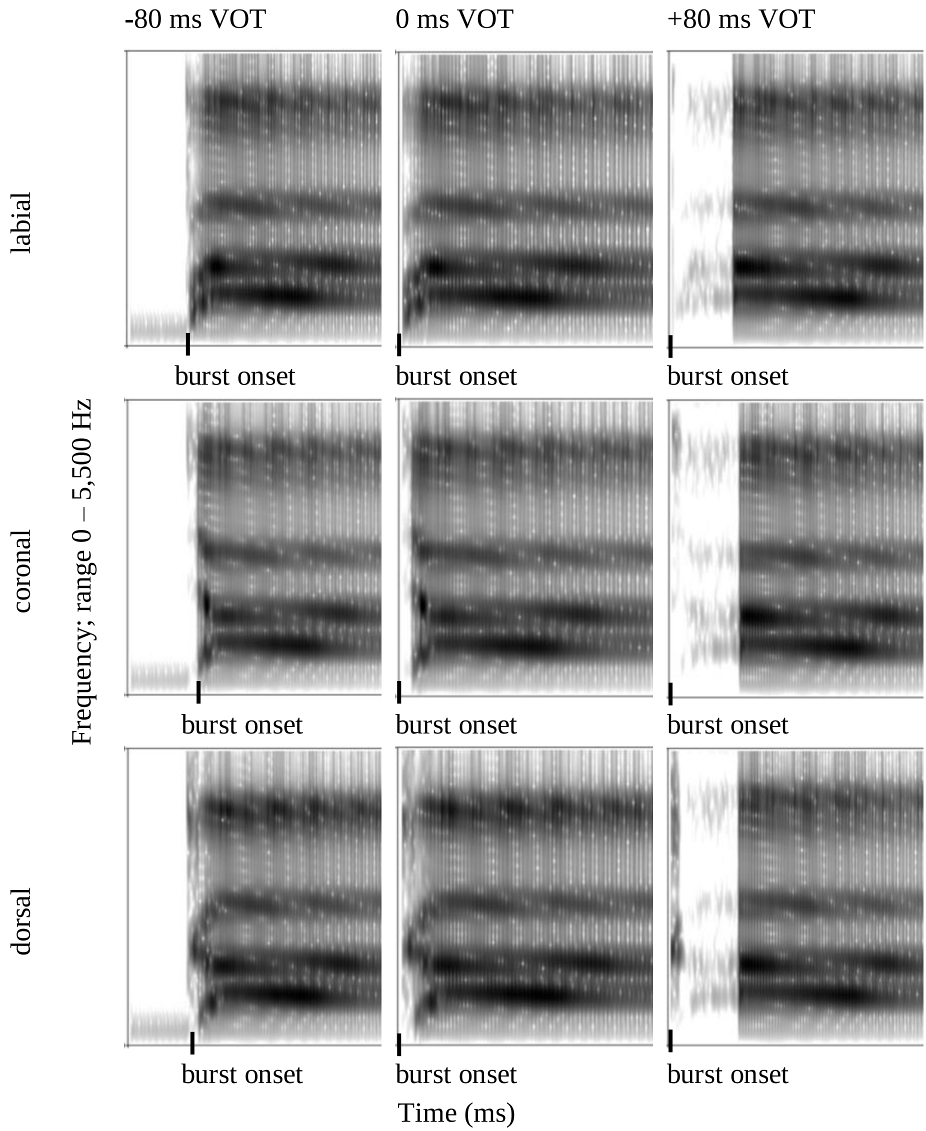


Figure 1. Formant transitions for labial, coronal, and dorsal plosives with VOT -80 ms, 0 ms, and +80 ms. Tick marks indicate the location of the burst onset.

computer mouse. Testing took place in a quiet room at Radboud University (Dutch native speakers) and the University of Cologne (German native speakers).

The experiment started with a practice block in which each end stimulus (i.e., VOT of -80 ms and +80 ms) was presented twice in random order. When accuracy on these 12 items was below 80% (Nittrouer & Miller, 1997), the practice block was repeated once[9]. After the critical threshold of 80% was passed, the test phase started.

In the test phase, 99 stimuli (i.e., 33 stimuli per place of articulation) with four repetitions each were presented in fully randomized order, with the exception that the same stimulus could not occur twice in a row. Breaks were offered after each 66 stimuli. The total duration of the experiment was approximately 15 minutes.

2.3 Results

Initial data screening revealed that in 455 (4%) of the 12,672 trials, the participants' response did not match the stimulus' place of articulation (Dutch: 328/6,336 trials=5%; German: 127/6,336 trials=2%). As can be seen in the confusion matrix displayed in Table 2, the largest number of miscategorized stimuli was prevoiced coronal stimuli, which the Dutch native listeners categorized as <ga> in 138 cases. In addition, the Dutch native listeners perceived prevoiced dorsal stimuli as <ba> in 61 cases. The German native listeners predominantly miscategorized aspirated coronal stimuli as <pa> (63 cases). Both groups were highly accurate identifying labial stimuli. Overall, the Dutch native listeners miscategorized prevoiced dorsal stimuli in 9.57% of the trials. The trials in which the participants' responses did not match the stimulus' place of articulation were excluded from the analysis.

Two mixed effects logistic regressions were performed in R (R Core Team, 2013), adopting an α-level of .05. The first analysis addressed the participants' perception of voicing in labial and coronal VOT continua, which are native to both Dutch and German. The second analysis addressed Dutch and German native speakers' perception of voicing in dorsal VOT continua and was conducted separately as the Dutch phonology does not encompass the 'voiced' dorsal plosive /g/.

[9] One native speaker of Dutch and three native speakers of German passed the 80% correct threshold after the second practice round. One native speaker of German did not pass the critical threshold after the second practice round. The participant nevertheless completed the experiment, but the data were excluded from analysis.

Table 2. Confusion matrix: Dutch and German native listeners.

	Dutch native listeners						German native listeners					
	<ba>	<da>	<ga>	<pa>	<ta>	<ka>	<ba>	<da>	<ga>	<pa>	<ta>	<ka>
labial prevoiced		2	1		1	0		1	1		0	1
coronal prevoiced	16		138	1		6	6		4	0		0
dorsal prevoiced	61	28		8	1		3	8		1	1	
labial 0 ms		0	0		0	0		0	0		0	0
coronal 0 ms	0		3	0		3	0		1	0		0
dorsal 0 ms	2	1		0	0		0	0		0	0	
labial aspirated		1	2		3	12		8	0		6	8
coronal aspirated	0		2	0		28	0		1	63		10
dorsal aspirated	3	0		0	3		3	1		0	0	

2.3.1 Perception of voicing in labial and coronal plosives

The participants' voicing categorization (0=voiced, 1=voiceless) was modeled with the fixed effects VOT (-80 ms to +80 ms), Language (Dutch=-1, German=1) and Place of Articulation (labial=-1, coronal=1). A three-way interaction term between VOT, Language, and Place of Articulation, as well as lower-level interaction terms were included. As random effects, the model included by-participant intercepts, and by-participant random slopes for VOT and Place of Articulation.

As expected, the results revealed a significant positive main effect for VOT (β=0.339, *SE*=0.090, *z*=3.765, *p*<.001), indicating that the participants categorized stimuli with longer VOT as 'voiceless'. In addition, a significant negative main effect for Language (β=-1.217, *SE*=0.221, *z*=-5.507, *p*<.001) shows that native speakers of Dutch (coded as -1) perceived more VOT stimuli as 'voiceless' than the native speakers of German (coded as 1). This means that the perceptual voicing boundary for the German native speakers is located on higher VOT values compared to the Dutch native speakers. The model detected significant negative two-way interactions between VOT and Place of Articulation (β=-0.026, *SE*=0.012, *z*=-2.212, *p*=.0.027), Language and Place of Articulation (β=-0.357, *SE*=0.163, *z*=-2.192, *p*=.028), and a significant negative three-way interaction between VOT, Language and Place of Articulation (β=-0.051, *SE*=0.012, *z*=-4.350, *p*<.001). The results of the main analysis are provided in Table 3 and visualized in Figure 2.

Separate by-group and by-place-of-articulation analyses were conducted to explore these interactions. Both by-group models showed a significant positive main effect for VOT, which confirms that both groups relied on VOT in their voicing perception (Dutch native speakers: β=0.427, *SE*=0.077, *z*=5.551, *p*<.001; German native speakers: β=0.442, *SE*=0.051, *z*=8.756, *p*<.001). The model on the German native speakers further revealed a significant negative interaction between VOT and Place of Articulation (β=-0.109, *SE*=0.025, *z*=-4.303, *p*<.001), which shows that the German native speakers had a steeper category boundary for coronal (coded as -1) than for labial (coded as 1) stimuli.

The model on the Dutch native speakers additionally revealed a significant positive main effect for Place of Articulation (β=0.470, *SE*=0.161, *z*=2.919, *p*=.004), suggesting that they perceived more labial (coded as +1) than coronal (coded as -1) stimuli as 'voiceless'. This means that the Dutch native speakers' category boundary for labial stimuli was located on lower VOT values than their category boundary for coronal stimuli. The model also revealed a significant positive interaction between VOT and Place of Articulation (β=0.026, *SE*=0.013, *z*=2.031, *p*=.042), suggesting that in opposition to the German native speakers, the Dutch native speakers had a steeper category boundary for labial (coded as 1) than for coronal (coded as -1) stimuli.

The by-place-of-articulation models both showed a significant positive main effect for VOT, which confirms that both groups relied on VOT in their voicing perception in both labial and coronal VOT continua (labial: β=0.445, SE=0.109, z=4.088, p<.001; coronal: β=0.045, SE=0.182, z=2.503, p=.012). The model on labial VOT continua further revealed a significant negative main effect for Language (β=-1.868, SE=0.212, z=-8.799, p<.001), which shows that the Dutch native speakers (coded as -1) perceived more labial stimuli as 'voiceless' than the German native speakers (coded as 1). Similarly, the model on coronal VOT continua showed a significant negative main effect for Language (β=-1.142, SE=0.214, z=-5.347, p<.001), suggesting that the Dutch native speakers (coded as -1) perceived more coronal stimuli as 'voiceless' than the German native speakers (coded as 1). In sum, the results of the by-place-of-articulation models confirm that the perceptual voicing category boundaries of the Dutch native listeners are located on lower VOT values for both labial and coronal VOT continua than the category boundaries of the German native speakers.

Table 3. Results of the main analysis on labial and coronal VOT continua.

	β	SE	z	p
Intercept	-1.031	0.221	-4.668	<.001
VOT	0.339	0.090	3.765	<.001
Language	-1.217	0.221	-5.508	<.001
Place of Articulation	-0.015	0.163	-0.091	>.250
VOT x Language	-0.089	0.090	-0.986	>.250
VOT x Place of Articulation	-0.026	0.012	-2.212	.027
Language x Place of Articulation	-0.358	0.163	-2.192	.028
VOT x Language x Place of Articulation	-0.051	0.012	-4.350	<.001

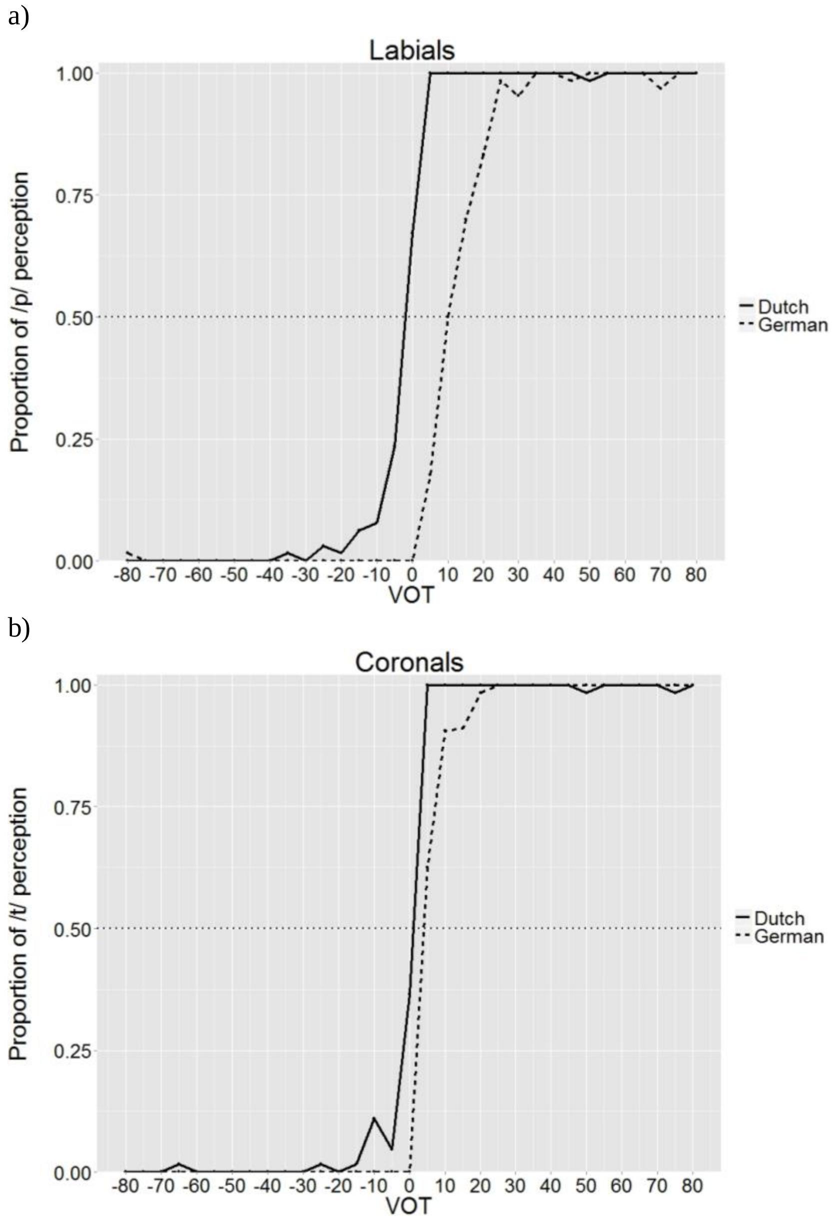


Figure 2. The proportion of 'voiceless' responses by the Dutch native speakers (solid line) and German native speakers (dashed line) to labial (top) and coronal VOT continua (bottom; over participants).

3

2.3.2 Perception of voicing in dorsal plosives

The participants' voicing categorization for dorsal VOT continua (0=voiced, 1=voiceless) was modeled with the fixed effects VOT (-80 ms to +80 ms) and Language (Dutch=-1, German=1), including a two-way interaction term. As random effects, the model included by-participant intercepts and random slopes for VOT.

The model revealed a significant positive main effect for VOT (β=0.35, *SE*=0.122, *z*=2.877, *p*<.004), which shows that both groups used VOT as a cue to voicing perception. In addition, the model detected a significant negative main effect for Language (β=-1.231, *SE*=0.142, *z*=-8.639, *p*<.001), suggesting that Dutch listeners' category boundary for dorsal VOT continua was located on lower VOT values compared to German listeners. The results are displayed in Table 4 and visualized in Figure 3.

Table 4. Results of the analysis on dorsal VOT continua.

	β	*SE*	*z*	*p*
Intercept	-0.923	0.142	-6.480	<.001
VOT	0.351	0.122	2.877	.004
Language	-1.231	0.142	-8.639	<.001
VOT x Language	0.120	0.122	0.987	>.250

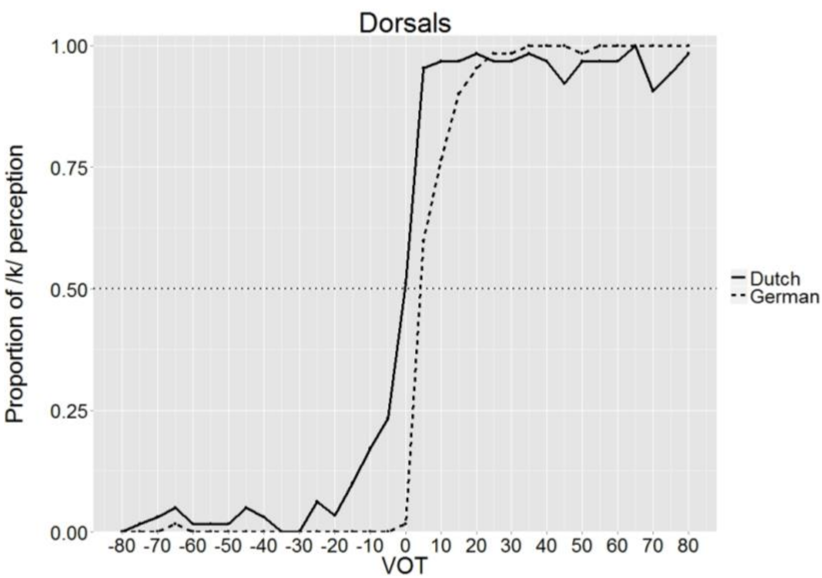


Figure 3. The proportion of 'voiceless' responses by the Dutch native speakers (solid line) and German native speakers (dashed line) to dorsal VOT continua (over participants).

2.3.3 Category boundaries

To determine the specific category boundaries for each group and place of articulation, by-group and by-place-of-articulation logistic regression analyses were conducted. The participants' voicing categorization (0=voiced, 1=voiceless) was modeled with the fixed effect VOT (-80 ms to +80 ms). As random effects, the model included by-participant intercepts and random slopes for VOT.

The intercept and slope values calculated by the logistic regressions (Table 5) entered the log odds function $\hat{y} = b + mx$, where $\hat{y}$ corresponds to the 0.5 point on the y-axis, *b* to the intercept, *m* to the slope, and *x* to the VOT value in question. Solving the equation for *x* provided the following category boundaries for Dutch native listeners: -0.70 ms for labials, +1.19 ms for coronals and +1.16 ms for dorsals. The category boundaries for German native listeners were: +10.05 ms for labials, +4.68 ms for coronals and +5.46 ms for dorsals.

3

Table 5. Results of the by-group and by-place-of-articulation analyses.

		Dutch				*German*			
		β	*SE*	*z*	*p*	*β*	*SE*	*z*	*p*
Labial	Intercept	0.890	0.339	2.624	.009	-2.907	0.308	-9.424	<.001
	VOT	0.560	0.087	6.494	<.001	0.339	0.080	4.238	<.001
Coronal	Intercept	0.024	0.285	0.085	>.250	-3.955	0.618	-6.403	<.001
	VOT	0.400	0.115	3.476	<.001	0.952	0.152	6.251	<.001
Dorsal	Intercept	0.307	0.097	3.148	.002	-1.866	0.318	-5.869	<.001
	VOT	0.166	0.028	5.943	<.001	0.433	0.067	6.477	<.001

2.4 Discussion

Dutch and German adults are known to differ in their VOT production, and this study showed that they also differ in their mapping of VOT onto voicing categories (Stoehr, Benders, van Hell & Fikkert, 2017a / Chapter 5). Dutch listeners resemble for example Spanish and French native speakers in their perception of VOT (respectively: Abramson & Lisker, 1973; Hazan & Boulakia, 1993), while German listeners perceive VOT similarly to English native listeners (Gonzales & Lotto, 2013; Lisker & Abramson, 1970). Importantly, this study showed that Dutch and German native listeners can make voicing judgements based on VOT alone.

For both groups, the specific category boundaries differed depending on the consonantal place of articulation. For Dutch native listeners, the category boundary was at higher VOT values when the consonantal place of articulation was more to the back of the mouth, resulting in a category shift on lowest VOT values for labial plosives. This finding is in line with previous research on VOT perception (Abramson & Lisker, 1973; Lisker & Abramson, 1970) and VOT production (Lisker & Abramson, 1964). Surprisingly, the German native listeners showed a different pattern in which the category boundary for labial stimuli occurred on higher VOT values than the category boundary for coronal and dorsal stimuli. This outcome may be related to the distribution of ‘voiceless’ versus ‘voiced’ plosives in German and differences in burst duration between the consonantal places of articulation, as explained below.

Based on CELEX counts (Baayen, Piepenbrock & van Rijn, 1993, accessed via the Reelex interface, Reetz, 2014), coronal and dorsal plosives follow a similar ‘voiced-voiceless’ distribution in German: word-initially, there are more tokens starting with ‘voiceless’ coronal plosives (66% of the tokens) and ‘voiceless’ dorsal plosives (70% of the tokens) compared to tokens starting with the ‘voiced’ counterpart plosive (34% of the coronal tokens and 30% of the dorsal tokens). This uneven distribution between ‘voiced’ and ‘voiceless’ plosives could create a bias towards labeling coronal and dorsal stimuli with ambiguous VOT as ‘voiceless’, which would result in a relatively early category shift in coronal and dorsal VOT continua. The ‘voiced-voiceless’ distribution is reversed for labial plosives in word-initial position. Only 45% of the tokens start with a ‘voiceless’ labial plosive. Due to these distributional differences, the finding that the category boundary for labial stimuli was located on higher VOT values than the boundaries for coronal and dorsal stimuli may be related to a perceptual bias towards /t/ and /k/ in coronal and dorsal VOT continua.

The overall later category shift in the labial VOT continuum can also be explained by differences between burst duration in labial plosives on the one hand and coronal and dorsal plosives on the other hand. For coronal and dorsal plosives, the burst duration was set to 15 ms while it only lasted 5 ms in labial plosives. It is possible that labial plosives with the shorter burst duration required some additional aspiration to be perceived as ‘voiceless’. The differences in burst duration may not play a crucial role for Dutch listeners as their main cue to voicing appeared to be the presence versus absence of prevoicing.

The second aim of Experiment I was to test whether Dutch native listeners generalize the voicing contrast to the dorsal place of articulation. For this reason, the Dutch listeners were presented with non-native speech sounds, in the form of prevoiced dorsal plosives. They mapped these onto the non-native segment /g/ in more than 90% of the trials. In line with labial and coronal stimuli, the Dutch listeners' perception of dorsal plosives changed rapidly from 'voiced' to 'voiceless' when no prevoicing was present.

Categorization of prevoiced dorsal plosives as /g/ is in line with PAM's perception as uncategorized speech sound. Miscategorization as /b/ or /d/, which would have been in line with PAM's assimilation to native speech sounds, occurred in less than 10% of the trials. However, we cannot rule out that the Dutch native speakers' categorization of prevoiced dorsal stimuli as /g/ is based on generalization in their L2 phonology rather than based on their native Dutch phonology. Although L2 experience and usage of the Dutch participants was minimal, all of them had some experience with English given that English language classes are obligatory during high school in the Netherlands.

Moreover, Experiment I provided practical insight and confirmed that the newly synthesized VOT continua are suitable to test Dutch and German native speakers' voicing perception. Both groups virtually never miscategorized labial stimuli for a different place of articulation, suggesting that the labial stimuli sounded most natural. In the following, we move on to Experiment II in which the voicing perception of Dutch-German bilingual and monolingual children is tested.

3 Experiment II: Children

3.1 Research questions

The present experiment is designed to answer the following three research questions:

1) Do Dutch-German simultaneous bilingual preschoolers have language-specific category boundaries between 'voiced' and 'voiceless' plosives along the VOT continuum?
2) Are language-specific voicing boundary locations and slopes associated with a child's current heritage language exposure?
3) Do bilingual children differ from their monolingual peers in the location and slope of their category boundaries between 'voiced' and 'voiceless' plosives along the VOT continuum?

3.2 Method

3.2.1 Participants

Fifty children participated in the experiment: 19 bilinguals (M_{age}=4;10 years, SD_{age}=9 months, range=3;9–5;11 years, 11 female), 16 Dutch monolinguals (M_{age}=5;0 years, SD_{age}=10 months, range=3;6–6;0 years, 9 female), and 15 German monolinguals (M_{age}=4;7 years, SD_{age}=11 months, range=3;6–6;0 years, 12 female). Of the 19 bilingual children, 11 children contributed data of the Dutch and German test sessions (see below for exclusion criteria). Three children only contributed data of the Dutch session and five children only contributed data of the German session. The 14 bilingual children who contributed Dutch perception data had an average age of 4;8 years (SD_{age}=9 months, range= 3;9–5;8 years, 9 female) and the 16 children who contributed German perception data had an average age of 4;11 years (SD_{age}=10 months, range= 3;9–5;11 years, 10 female).

Another 47 children were tested, but excluded from the analyses because they either did not meet the inclusion criteria[10] (4 bilinguals, 4 monolinguals), refused to do the task (3 bilinguals, 2 German monolinguals), exhibited a side bias, that is they responded based on the side of presentation rather than based on the auditory stimuli

[10] Three bilingual children were exposed to a third language and one bilingual child's onset of bilingualism was after the first year of life; three Dutch monolingual children and one German monolingual child were exposed to a non-native speaker of their native language in their immediate social environment.

(4 bilinguals, 9 Dutch monolinguals, 7 German monolinguals), or responded inaccurately to the least ambiguous stimuli, henceforth called end stimuli (3 bilinguals, 5 Dutch monolinguals, 6 German monolinguals). Determination of side biases and endpoint accuracy are addressed in detail below.

3.2.2 Materials

The experiment was designed as an animated video game to create a child-friendly experimental setting and to enhance children's attention during the experiment. The game consisted of video clips showing a cartoon dinosaur mother and two dinosaur babies or a panda mother and two panda babies. The mother character was in the center of the screen, with one baby character to her left, and one to her right (Figure 5). All video materials were created with Adobe Photoshop and Flash CS 5, and were a subset of those created and used by Zhou (2015).

The auditory stimuli were the same as in Experiment I. To reduce the duration of the experiment, a narrower VOT range of -35 ms to +40 ms was used. This range comprises eight stimuli without aspiration (-35 ms to 0 ms) and eight stimuli with aspiration (+5 ms to +40 ms). A similar VOT range had previously been used in a study on adult bilinguals' voicing perception (Gonzales & Lotto, 2013).

During practice, four triads of monosyllabic pseudowords were presented (Table 6). These pseudowords were phonotactically legal and encompassed Dutch-specific phonemes in the Dutch task, and German-specific phonemes in the German task.

Three female native speakers of Dutch and three female native speakers of German were recorded in a child-directed register for the pseudoword practice stimuli as well as for the instructions and encouraging phrases played before and during the experiment. Recordings were made in a sound-attenuated booth using Adobe Audition software (sampling rate: 44.1 kHz, channels: stereo, resolution: 16-bit).

3.2.3 Design and procedure

The perception experiment was embedded in a testing session in which production data of the children (Stoehr et al., 2017b / Chapter 2) and parents (Stoehr et al., 2017a / Chapter 5) were collected. The testing session took place in the children's homes. Prior to the perception experiment, the children completed a speech production task in the target language and interacted with the experimenter in the

Table 6. Dutch and German pseudowords used in the practice blocks.

		Dutch			German		
Block	Trial	Mother	Baby (left)	Baby (right)	Mother	Baby (left)	Baby (right)
1	1	[zœym]	[vaːm]	[zœym]	[ʀɔp͡f]	[mʏp͡f]	[ʀɔp͡f]
	2	[zœym]	[zaːm]	[zœym]	[ʀɔp͡f]	[ʀʏp͡f]	[ʀɔp͡f]
	3	[zœym]	[zœym]	[vœym]	[ʀɔp͡f]	[ʀɔp͡f]	[mɔp͡f]
	4	[miːɫ]	[miːɫ]	[ʀoːɫ]	[ʃɪl͡ts]	[ʃɪl͡ts]	[jal͡ts]
	5	[miːɫ]	[moːɫ]	[miːɫ]	[ʃɪl͡ts]	[ʃal͡ts]	[ʃɪl͡ts]
	6	[miːɫ]	[miːɫ]	[ʀiːɫ]	[ʃɪl͡ts]	[ʃɪl͡ts]	[jɪl͡ts]
2	1	[jɪɫ]	[jɪɫ]	[fœyɫ]	[t͡saɪf]	[nif]	[t͡saɪf]
	2	[jɪɫ]	[jœyɫ]	[jɪɫ]	[t͡saɪf]	[t͡saɪf]	[t͡sif]
	3	[jɪɫ]	[jɪɫ]	[fɪɫ]	[t͡saɪf]	[naɪf]	[t͡saɪf]
	4	[xɑɫf]	[nɛɫf]	[xɑɫf]	[fɔʏn͡ts]	[fɔʏn͡ts]	[t͡saʊn͡ts]
	5	[xɑɫf]	[xɑɫf]	[xɛɫf]	[fɔʏn͡ts]	[faʊn͡ts]	[fɔʏn͡ts]
	6	[xɑɫf]	[nɑɫf]	[xɑɫf]	[fɔʏn͡ts]	[fɔʏn͡ts]	[t͡sɔʏn͡ts]

target language for approximately 30 minutes. During this period, the parents were also instructed to communicate with the child exclusively in the target language. The experimental sessions were always conducted by a native speaker of the target language.

Parents gave informed consent for their child to participate in the study. Either prior to testing or during testing, the parents completed a language history questionnaire, which was based on the BiLEC (Unsworth, 2013) for bilingual children and custom-made for the monolingual children.

A speech sound categorization task programmed in Presentation software (Neurobehavioral systems, version 14.7) was administered on a Hewlett Packard EliteBook 8540P with a resolution of 1366 x 768 pixels. The experimental script created and used by Zhou (2015) was modified for the present experiment. Auditory stimuli were presented over Sennheiser HD 280 64 ohm headphones, and responses were recorded via a two-button response box (MPI Dual Button box: Serial port via USB, Baudrate-38400, Data-8bit, and StopBit-1). The experiment consisted of two or three practice blocks (depending on accuracy, as explained below), and nine experimental blocks. Table 7 provides an outline of the experiment.

Table 7. Outline of the experiment.

	Block number	Condition	Number of trials	Feedback
	1	Left and right (non-linguistic)	4	yes
Practice	2	Pseudowords	6	yes
	(3	Pseudowords	6	yes)
	1	Places of articulation	6	no
	2	VOT endpoints	6	no
	3	VOT labial plosives	2 x 8	no
Test	4	VOT coronal plosives	2 x 8	no
	5	VOT labial plosives	2 x 8	no
	6	VOT coronal plosives	2 x 8	no
	7	VOT dorsal plosives	2 x 8	no
	8	VOT dorsal plosives	2 x 8	no
	9	Places of articulation	6	no

The speech sound categorization task was set up in an XAB format, which has previously been used to investigate categorical speech perception in children aged three to eleven years (Brasileiro, 2009; Giezen, 2011; Zhou, 2015). In the XAB task, the children had to decide whether the first stimulus (X) was more similar to the second (A) or third (B) stimulus. Throughout the practice and test XAB-trials, the X stimulus was provided by the dinosaur mother (practice & experimental blocks 1–6) or the panda mother (experimental blocks 7–9), and was played 1,000 ms after the onset of the trial. The A stimulus was provided through the baby figure left to the mother and started 1,500 ms after the offset of the X stimulus. The A and B stimuli had an offset to onset inter-stimulus interval of 1,000 ms to target phonological processing (Werker & Logan, 1985). The children gave their response after the offset of stimulus B through a two-button response box, on which the left button corresponded to the baby figure left of the mother and the right button corresponded to the baby figure right of the mother (see Figure 5). After each response, a third baby figure, which never spoke and was displayed in a separate section on the far right of the screen, jumped one step up. After the child responded to six trials (practice blocks, experimental blocks 1, 2 and 9) or eight trials (experimental blocks 3–8), the baby figure on the far right reached the gift at the top step, and its content was revealed.

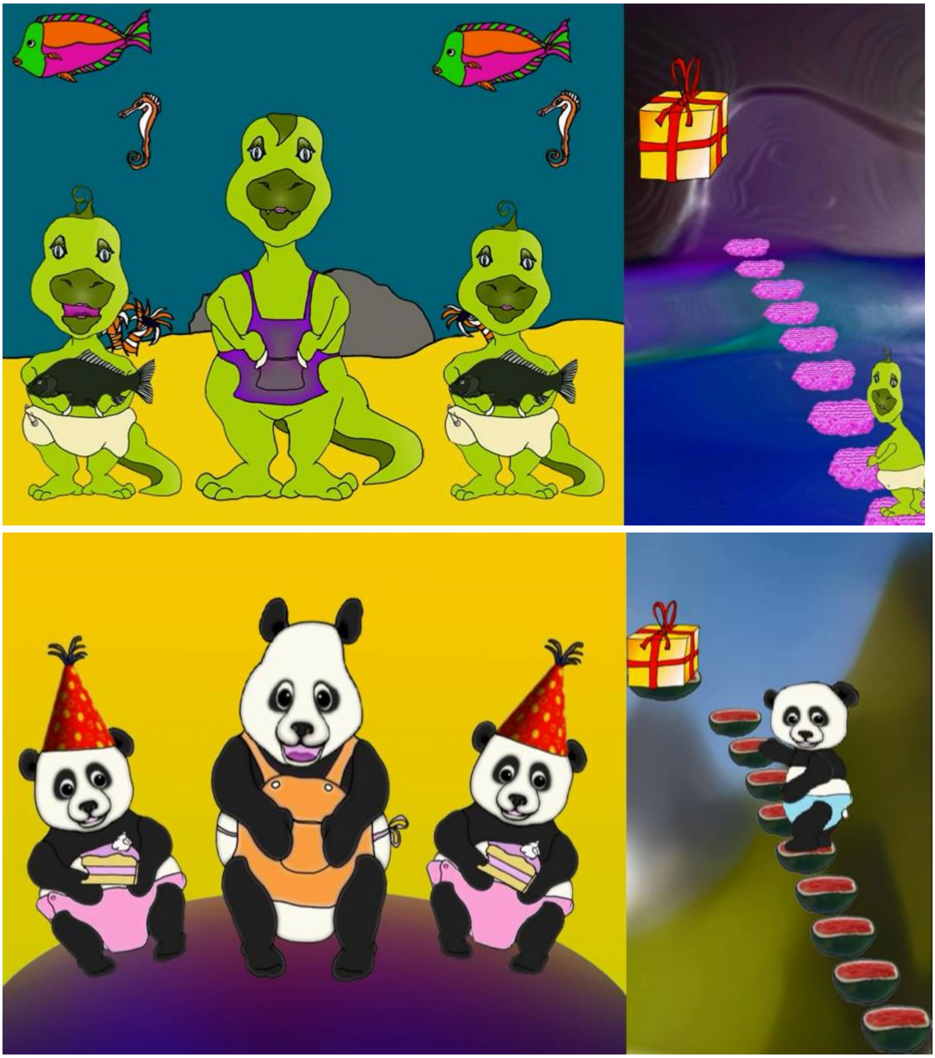


Figure 5. Example videos during XAB trials.

Practice blocks. Two practice units preceded the experiment to familiarize the children with the task. During the first practice block, the children were trained to give responses through the two-button response box. The first block consisted of four trials, in which one of the two baby dinosaurs jumped up and down and verbally encouraged the child to press her button. When the child gave the correct response, the dinosaur mother provided positive feedback and the next trial started. When the child gave an incorrect response, the baby dinosaur continued to jump up and down.

The dinosaur mother pointed at the baby dinosaur and instructed the child to press the corresponding button.

The second practice block used the XAB paradigm with the pseudoword stimuli. Per trial, the X stimulus pseudoword was presented through the dinosaur mother. The A and B stimuli were provided through the two baby dinosaurs, who repeated the pseudoword either matching or mismatching the dinosaur mother's production. Each X stimulus pseudoword was presented in three trials, first paired with a correct response and a mispronunciation in which the vowel and the onset consonant differed. In the second trial, the mispronunciation only differed from the X stimulus in the vowel. In the third trial, the mispronunciation only differed from the X stimulus in the onset consonant. Two pseudoword triads were presented per block, resulting in a total of six trials. After each response, feedback was provided through the dinosaur mother. A correct response was verbally praised and animated fireworks appeared in the background. When the response was incorrect, the dinosaur mother pointed at the baby dinosaur who gave the correct response and verbally encouraged the child to press the corresponding button. Only when the child gave the correct response, the next trial started. When the child answered at least four out of six trials correctly on the first attempt of this practice block, the experiment started. If accuracy was lower, an additional practice block started, in which two new triads of pseudowords were presented.

Experimental blocks. In contrast to the practice blocks, no feedback was provided during the experimental blocks. In the first and the last (=9th) experimental blocks, children had to discriminate VOT endpoint stimuli in their place of articulation. In the second block, the children had to discriminate the VOT endpoint stimuli in voicing (i.e., -35 ms and +40 ms). To enhance a monolingual setting as well as to maintain a balanced presentation of prevoiced and aspirated items, no dorsal items were presented, as the 'voiced' dorsal plosive represent a phonological gap in Dutch. Instead, only labial (1x -35 ms, 2x +40 ms) and coronal (2x -35 ms, 1x +40 ms) stimuli were presented. The place of articulation discrimination blocks and the VOT endpoint discrimination block each contained six trials.

In the voicing categorization blocks (blocks 3–8), the X stimulus varied in VOT and was presented through the dinosaur mother (labial and coronal plosives) or panda mother (dorsal plosives). The A and B sounds were the 'voiced' (-35 ms) and 'voiceless' (+40 ms) endpoint stimuli and were presented through two baby figures, as depicted in Figure 5.

Each VOT categorization block contained 16 trials. After completion of the first eight trials, the gift on the right side of the screen was opened and the mother

3

figure praised the child and encouraged her to finish the second half of the block. The duration of the first half of the block was approximately two minutes, which means that the child was exposed to the test language through the mother figure approximately every two minutes. The first two X stimuli played at the first half of the block were always set to VOT values of -35 and +40 ms (counterbalanced), and the first two X stimuli of the second half of the block were always set to VOT values of -15 ms and +20 ms (counterbalanced). The order of the remaining stimuli was randomized. Stimuli in the critical VOT range with VOT values of -15 ms, -10 ms, -5 ms, 0 ms, +5 ms, +10 ms, +15 ms, and +20 ms were repeated three times per place of articulation throughout the experiment. Stimuli with VOT values of -35 ms, -30 ms, -25 ms, -20 ms, +25 ms, and + 30 ms were only presented once per place of articulation during the VOT categorization blocks.

The labial stimuli were presented in blocks 3 and 5, and the coronal stimuli were presented in blocks 4 and 6. To approach a monolingual setting for the bilinguals, and a natural setting for the monolingual Dutch children, the dorsal stimuli were presented at the end of the experiment in blocks 7 and 8 because ‘voiced’ dorsal plosives are non-native to Dutch. After each block, the experiment was paused and the experimenter rewarded the child with two stickers. The duration of the breaks between blocks depended on the individual child’s needs. At the beginning of each block, the mother figure greeted and motivated the child.

An obligatory extended break of at least ten minutes was taken before the dorsal stimuli in blocks 7 and 8 were presented. During this break, the children completed a second production task, which was designed as a game (Stoehr et al., 2017b / Chapter 2).

3.3 Results

3.3.1 Data screening

Each child’s proportion of left vs. right responses was calculated to identify children who exhibited a side bias. Side biases in behavioral testing with children of similar ages had previously been reported (Bergmann, Paulus & Fikkert, 2012; Simon & Fourcin, 1977). Children who responded with the same button in more than 70% of the relevant trials (i.e., voicing trials using labial and coronal stimuli) throughout the experiment were excluded from the analysis. Second, each remaining child’s response data were screened individually. When a child continuously pressed the same button throughout one block (six trials for place of articulation discrimination and VOT endpoint discrimination blocks; and eight trials for voicing categorization blocks), the block was removed from the analysis. In addition, trials with response times above 10,000 ms were excluded. This rather high cut-off point had previously been suggested for the specific experimental design and the investigated ages, as children frequently pointed to the screen, and only gave their response through the response box after the experimenter encouraged them to do so (Zhou, 2015).

Based on the remaining data, each child's accuracy in responding to the VOT end stimuli (-35 ms and +40 ms) was calculated. Accuracy cut-off values of 80% are commonly used in perception studies (Nittrouer & Miller, 1997), but due to the limited sample size and the small number of stimuli used in this study, the cut-off point was set to an endpoint identification accuracy of 50% or higher. To account for differences in endpoint identification, accuracy in identifying the VOT endpoints (-35 ms and +40 ms) was used as an independent variable in the statistical analyses (Giezen, 2011).

In the VOT categorization conditions, data cleaning based on the block-specific side biases and reaction times led to exclusion of 666 (16%) out of 4270 trials across participant groups (Dutch monolinguals: 98/1,120 trials [9%], German monolinguals: 229/1,050 trials [22%], bilinguals in Dutch: 204/980 trials [21%], bilinguals in German: 135/1,120 trials [12%]). In the place of articulation discrimination condition, 110 out of 732 trials (15%) were excluded across participant groups (Dutch monolinguals: 22/192 trials [11%], German monolinguals: 28/180 trials [16%], bilinguals in Dutch: 22/168 trials [13%], bilinguals in German: 32/192 trials [17%]).

After data cleaning, each child contributed on average two and a half trials for each labial and coronal stimulus per language in the critical VOT range of -15 ms to +20 ms. For each endpoint stimulus (i.e., -35 ms and +40 ms), each child contributed on average two trials per labial and coronal stimulus. For the stimuli with VOT values of -30, -25, -20, +25, +30, and +35, each child contributed on average 0.85 trials per stimulus and place of articulation.

The included children's accuracy in discriminating the VOT endpoint stimuli was on average 78% and did not differ across groups, $F(3, 57)=0.423$, $p>.250$ (Figure 6; Dutch monolinguals: M=80%, SD=15, range=50–100%; bilinguals in Dutch: M=79%, SD=13, range=55–100%; bilinguals in German: M=79%, SD=14, range=58–100%; German monolinguals: M=75%, SD=13, range=50–92%).

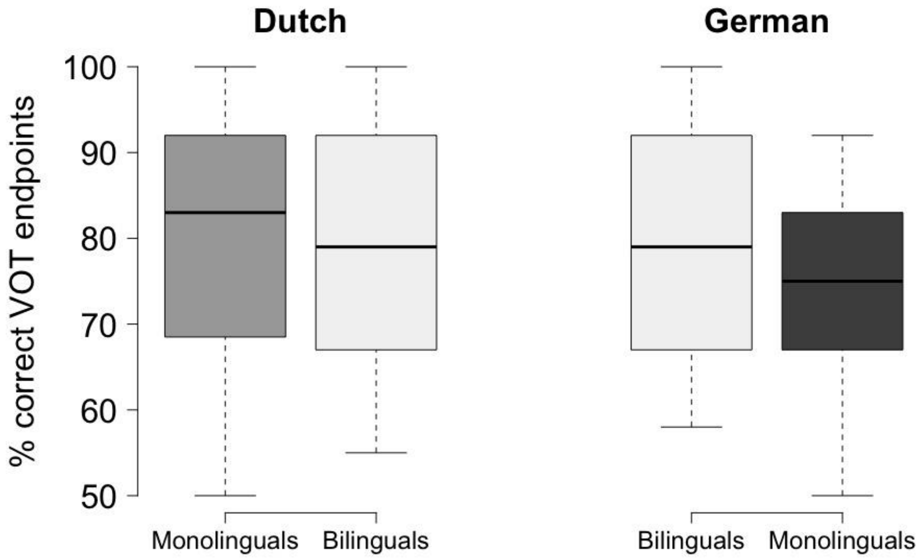


Figure 6. VOT endpoint discrimination accuracy by group and language (over participants).

3.3.2 Individual labelling patterns

Visual data inspection revealed that the children tested in this study had individual labeling patterns, which matched those described by Simon and Fourcin (1977). Categorical labeling was observed in 22 children (5 bilinguals in Dutch, 7 bilinguals in German, 6 Dutch monolinguals, and 4 German monolinguals). Progressive labeling was observed in 7 children (1 bilingual in Dutch, 3 bilinguals in German, 2 Dutch monolinguals, and 1 German monolingual). One monolingual German child's labeling function approached a flat line. The most frequently observed pattern was scattered labeling which was observed in 29 children (7 bilinguals in Dutch, 5 bilinguals in German, 8 Dutch monolinguals, and 9 German monolinguals). An interesting pattern emerges when the age of the children is taken into account for each labeling pattern (Figure 7). The monolinguals show a developmental trend from scattered labeling over progressive labeling to categorical labeling with increasing age. No such age trend was observed in the bilinguals.

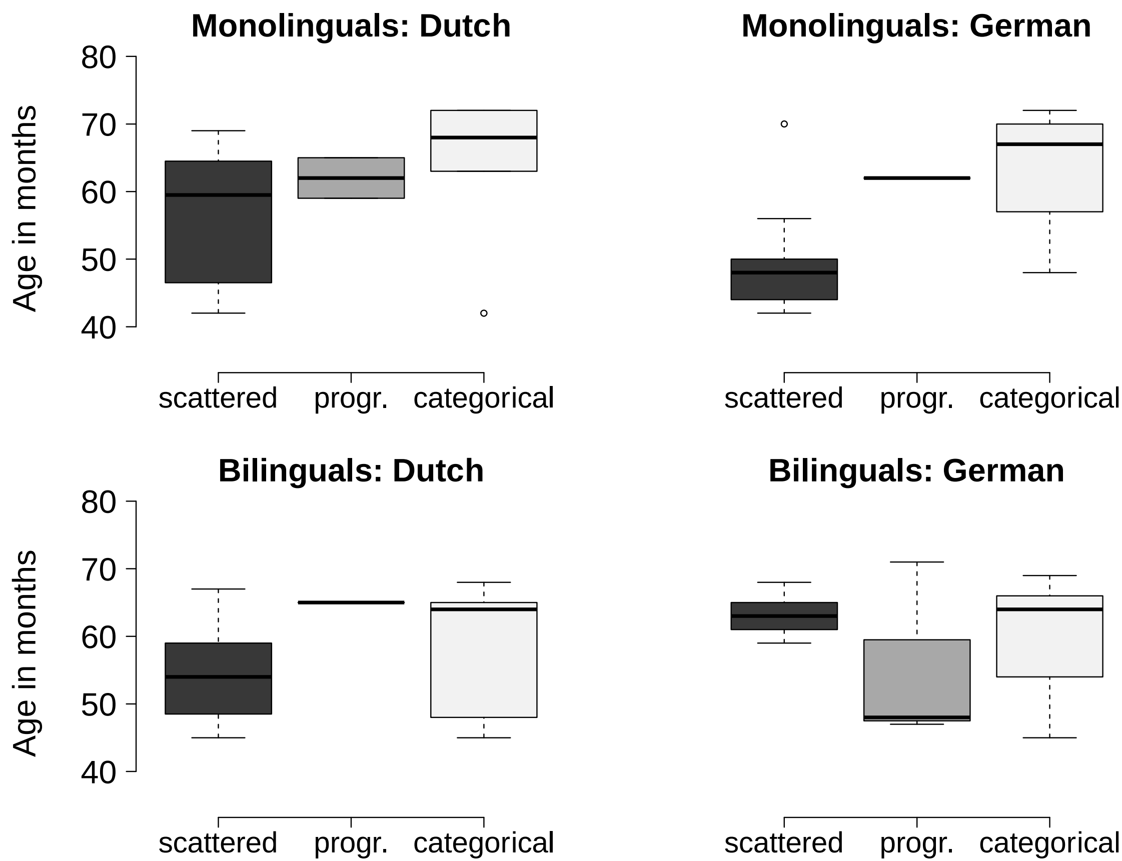


Figure 7. Age and labeling functions by groups over children

3.3.3 Mixed effects logistic regression analysis

Mixed effects logistic regressions were performed in R (R Core Team, 2013), adopting an α-level of .05. We initially planned to include all places of articulation in the analysis of the child data, but decided to base the statistical analyses only on labial stimuli for the following three reasons: 1) The numeric difference between the Dutch and German adults' perceptual boundaries was most prominent for labials (Dutch adults: -0.70 ms; German adults: +10.05 ms; difference: 10.75 ms), but differed only minimally for coronals (Dutch adults: +1.19 ms; German adults: +4.68 ms; difference: 3.49 ms) and dorsals (Dutch adults: +1.16 ms; German adults: +5.46 ms; difference: 3.52); 2) the adult experiment revealed that labial stimuli were miscategorized less frequently (1.11%) than coronal stimuli (6.67%) and dorsal (2.98%) stimuli, which suggests that the synthesized labial stimuli sounded most natural to Dutch and German native listeners; 3) labial stimuli were presented early in the experiment and always after short breaks, suggesting that the children were more focused when labial stimuli were presented. For the sake of completeness, the model reporting combined results of labial and coronal stimuli, that is, plosives that are native to both Dutch and German, can be found in Appendix 1, and the results of the model on dorsal plosives are available in Appendix 2.

The following analyses are based on the VOT stimuli between -30 ms and +35 ms. As the VOT endpoints -35 ms and +40 ms were used to compute the variable Endpoint Identification Accuracy, they were excluded from the analyses to ensure that the dependent variable was independent of this predictor variable.

3.3.4 Monolingual children

The monolingual children's voicing categorization (0=voiced, 1=voiceless) in labial VOT continua was modeled with the fixed effects VOT (-30 ms to +35 ms), Language (Dutch=-1, German=1), and Endpoint Identification Accuracy (centered around zero). Interaction terms between VOT and Language, and Endpoint Identification Accuracy were entered[11]. As random effects, the model included intercepts for child, as well as by-child random slopes for VOT.

[11] Due to the observed labeling patterns in the monolingual children, we ran one additional model including Age as a fixed effect and an interaction between Age and VOT. There was no significant effect for Age and no significant interaction between Age and VOT. To keep the models for monolinguals and bilinguals as comparable as possible, we only report the model excluding Age in this chapter.

The results revealed a significant positive main effect for VOT (β=0.052, *SE*=0.010, *z*=5.047, *p*<.001), showing that the children categorized stimuli with longer VOT as 'voiceless'. A significant negative main effect for Endpoint Identification Accuracy (β=-0.037, *SE*=0.011, *z*=-3.528, *p*<.001) and a significant positive interaction between VOT and Endpoint Identification Accuracy (β=0.003, *SE*=0.001, *z*=3.622, *p*<.001) suggest that children who identified the endpoint VOT stimuli more accurately used VOT more as a cue to voicing than children with lower endpoint identification accuracy. No other significant main effects or interactions were observed. These results are displayed in Table 8 and visualized in Figure 8.

Table 8. Results of the main analysis on the monolingual children's voicing perception in labial VOT continua.

	β	*SE*	*z*	*p*
Intercept	-0.348	0.147	-2.369	.018
VOT	0.052	0.010	5.047	<.001
Language	-0.116	0.148	-0.786	>.250
Endpoint Accuracy	-0.037	0.011	-3.528	<.001
VOT x Language	-0.005	0.010	-0.471	>.250
VOT x Endpoint Accuracy	0.003	0.001	3.622	<.001

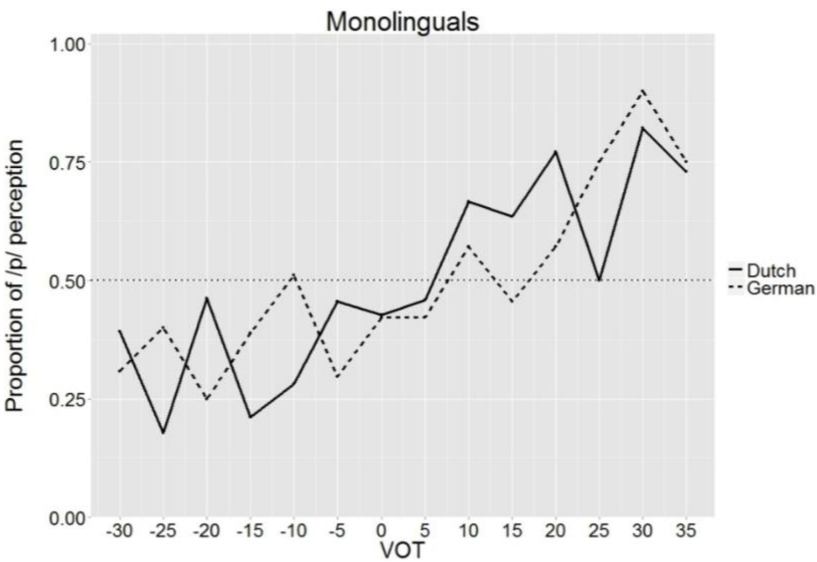


Figure 8. The proportion of 'voiceless' responses by the Dutch monolingual children (solid line) and German monolingual children (dashed line) to labial VOT continua (over participants).

3.3.5 Bilingual children

The bilingual children's voicing categorization (0=voiced, 1=voiceless) was modeled with the fixed effects VOT (-30 ms to +35 ms), Language (Dutch=-1, German=1), Percent of Exposure to German at the time of testing (centered around zero; henceforth Exposure) and Endpoint Identification Accuracy (centered around zero). Interaction terms between VOT and Language, Language and Exposure, and VOT and Endpoint Identification Accuracy were entered. As random effects, the model included intercepts for child, as well as by-child random slopes for Language and VOT.

The results revealed a significant positive main effect for VOT (β=0.039, *SE*=0.009, *z*=4.337, *p*<.001), showing that the children categorized stimuli with longer VOT as 'voiceless'. A significant negative main effect for Language (β=-0.195, *SE*=0.095, *z*=-2.047, *p*=.041) suggests that the bilingual children perceived more labial stimuli as 'voiceless' in the Dutch testing session compared to the German testing session. This means that their category boundary is located on lower VOT values in Dutch than in German. Lastly, a significant negative main effect for Endpoint Identification Accuracy (β=-0.021, *SE*=0.008, *z*=-2.544, *p*=.011) and a significant positive interaction between Endpoint Identification Accuracy and VOT (β=0.002, *SE*=0.001, *z*=2.870, *p*=.004) suggest that children who identified the endpoint stimuli more accurately used VOT more as a cue to voicing than children with lower accuracy identifying the VOT endpoint stimuli. No other significant main effects or interactions were observed. The results are displayed in Table 9 and visualized in Figure 9.

Table 9. Results of the main analysis on the bilingual children's voicing perception in labial VOT continua.

	β	*SE*	*z*	*p*
Intercept	-0.442	0.128	-3.439	<.001
VOT	0.039	0.009	4.337	<.001
Language	-0.195	0.095	-2.047	.041
Exposure	0.007	0.007	1.071	>.250
Endpoint Accuracy	-0.021	0.008	-2.544	.011
VOT x Language	0.004	0.006	0.744	>.250
Language x Exposure	0.001	0.007	0.186	>.250
VOT x Endpoint Accuracy	0.002	0.001	2.870	.004

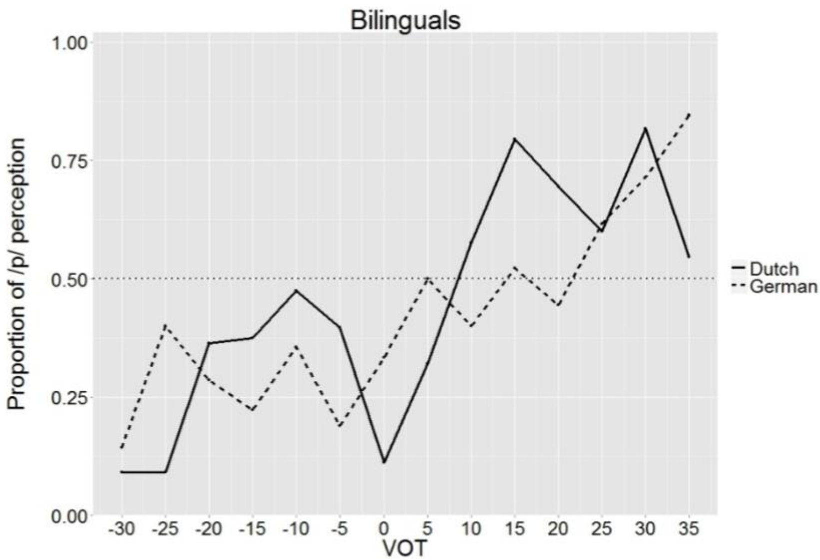


Figure 9. The proportion of 'voiceless' responses by the bilingual children in Dutch (solid line) and in German (dashed line) to labial VOT continua (over participants).

3.3.6 Bilingual children versus monolingual children

The children's voicing categorization (0=voiced, 1=voiceless) in labial VOT continua was modeled with the fixed effects VOT (-30 ms to +35 ms), Group (bilinguals=-1, monolinguals=1), and Endpoint Identification Accuracy (centered around zero). Interaction terms between VOT and Group, and Endpoint Identification Accuracy were entered. As random effects, the model included intercepts for child, as well as by-child random slopes for VOT. One model was run on the Dutch data and a second model was run on the German data.

Both models revealed a significant positive main effect for VOT (Dutch: β=0.055, *SE*=0.011, *z*=4.877, *p*<.001; German: β=0.040, *SE*=0.007, *z*=5.987, *p*<.001) and a significant negative main effect for Endpoint Identification Accuracy (Dutch: β=-0.027, *SE*=0.008, *z*=-3.455, *p*<.001; German: β=-0.037, *SE*=0.012, *z*=-3.233, *p*=.001). In addition, there was a significant positive interaction between VOT and Endpoint Identification Accuracy in both models (Dutch: β=0.003, *SE*=0.001, *z*=3.461, *p*<.001; German: β=0.002, *SE*=0.001, *z*=3.926, *p*<.001), suggesting that children who identified the endpoint stimuli more accurately used VOT more as a cue to voicing than children with lower accuracy identifying the VOT endpoint stimuli. No differences between bilinguals' and monolinguals' voicing perception were observed in either model (Dutch: β=0.073, *SE*=0.107, *z*=0.681, *p*>.250; German: *β*=0.067, *SE*=0.154, *z*=0.431, *p*>.250). Table 10 displays the results of both models.

Table 10. Results of the main analyses on voicing perception in labial VOT continua of monolingual and bilingual children in Dutch (left) and German (right).

	Dutch				German			
	β	*SE*	*z*	*p*	*β*	*SE*	*z*	*p*
Intercept	-0.360	0.107	-3.370	<.001	-0.497	0.152	-3.271	.001
VOT	0.055	0.011	4.877	<.001	0.040	0.007	5.987	<.001
Group	0.073	0.107	0.681	>.250	0.067	0.154	0.431	>.250
Endpoint Accuracy	-0.027	0.008	-3.455	<.001	-0.037	0.012	-3.233	.001
VOT x Group	0.007	0.011	0.065	>.250	0.001	0.007	0.209	>.250
VOT x Endpoint Acc.	0.003	0.001	3.461	<.001	0.002	0.001	3.926	<.001

3.3.7 Category boundaries

To determine the specific category boundary locations for each group and language, by-language mixed-effects logistic regression analyses were conducted for each group. The participants' voicing categorization (0=voiced, 1=voiceless) was modeled with the fixed effects VOT (-30 ms to +35 ms) and Endpoint Identification Accuracy (centered around zero) with an interaction term. The models on the bilinguals further included the fixed effect Exposure (centered around zero). As random effects, the model included intercepts for child, and by-child random slopes for VOT.

The intercept and slope values calculated by the logistic regressions (Table 11) entered the log odds function $\hat{y} = b + mx$, where $\hat{y}$ corresponds to the 0.5 point on the y-axis, *b* to the intercept, *m* to the slope, and *x* to the critical VOT value. The computed category boundary points were +11.73 ms for Dutch monolinguals, +22.35 ms for German monolinguals, +20.36 ms for the bilinguals in Dutch, and +26.76 ms for the bilinguals in German. Although the category shift of monolingual Dutch children appears to occur earlier than the category shift of German monolingual children, no significant differences were detected between the category boundary locations or slopes of the monolingual Dutch and German children. The bilingual children's category shift in labial plosives occurs 10 ms later in German than in Dutch.

Table 11. Results of the by-group and by-language analyses.

		Dutch				German			
		β	*SE*	*z*	*p*	β	*SE*	*z*	*p*
Monolinguals	Intercept	-0.333	0.160	-2.077	.038	-0.372	0.269	-1.382	.167
	VOT	0.071	0.021	3.425	<.001	0.039	0.010	3.993	<.001
	Endpoint Accuracy	-0.032	0.011	-2.876	.004	-0.051	0.021	-2.396	.017
	VOT x Endpoint Accuracy	0.004	0.001	2.748	.006	0.002	0.001	2.959	.003
Bilinguals	Intercept	-0.355	0.152	-2.335	.019	-0.624	0.156	-3.999	<.001
	VOT	0.042	0.011	3.951	<.001	0.042	0.010	4.213	<.001
	Endpoint Accuracy	-0.023	0.013	-1.843	.065	-0.029	0.012	-2.449	.014
	Exposure	0.007	0.011	0.569	>.250	0.008	0.009	0.906	>.250
	VOT x Endpoint Accuracy	0.002	0.001	2.255	.024	0.002	0.001	2.602	.009

4 Discussion

The results of the present speech perception study support that Dutch-German bilingual preschoolers have language-specific voicing systems. Bilingual preschoolers' language-specific voicing perception fills the gap between observations of preverbal bilingual infants' neural sensitivity to the voicing contrasts of both of their native languages (Ferjan Ramírez et al., 2016; Garcia-Sierra et al., 2011) and early bilingual adults' language-specific voicing perception (Elman et al., 1977; Garcia-Sierra et al., 2009; Gonzales & Lotto, 2013; Hazan & Boulakia, 1993).

The finding that bilingual preschoolers have language-specific voicing boundaries provides important novel insight into previous speech production data, which showed that bilingual children produced VOT language-specifically for 'voiceless' plosives, but not for 'voiced' plosives (Deuchar & Clark, 1996; Khattab, 2000; Johnson & Wilson, 2002 (the older child); Mayr & Siddika, 2016). This production pattern was even present in the recent VOT production results obtained from the same children as tested in the present perception study (Stoehr et al., 2017b / Chapter 2): the bilingual children prevoiced a statistically indistinguishable percentage of 'voiced' plosives in Dutch and German, and this percentage was intermediate to that of Dutch and German monolinguals. Stoehr et al. (2017b / Chapter 2) provided two possible interpretations of these results: the bilingual children either acquired a merged category for Dutch and German 'voiced' plosives, or they have two separate 'voiced' categories, which developed indistinguishably at the time of testing. The present VOT perception results provide the missing puzzle piece to understand these previous findings by showing that the bilingual children associate 'voiced' plosives with longer VOT values in German than in Dutch. This finding informs previous work on bilingual children's VOT production by suggesting that indistinguishable productions of 'voiced' plosives across two languages are nevertheless the result of two separate language-specific categories.

Language exposure was not found to play a crucial role in the bilingual children's voicing perception, although increasing heritage language exposure has previously been associated with more target-like production of 'voiceless' plosives in the same children (Stoehr et al., 2017b / Chapter 2). This suggests that the bilingual children perceived voicing in labial VOT continua language-specifically independent of their language exposure at the time of testing. Even the bilingual children with little exposure to either Dutch or German established language-specific phonological representations of /b/ and /p/ that are comparable to those of their peers with more balanced exposure to both languages, which is compatible with the assumption that perception precedes production in first language acquisition (e.g., Menyuk & Anderson, 1969).

Bilinguals and monolinguals did not differ detectably in their voicing perception in either language, although there were numerical differences in the category boundary location between the groups. Specifically, the voicing boundary was 8 ms later in bilinguals tested in Dutch (20 ms of VOT) than in Dutch monolinguals (12 ms of VOT), which is a difference in the expected direction towards a more German-like boundary location. However, in German, the bilinguals' voicing boundary was also 5 ms later in bilinguals (27 ms of VOT) than in monolinguals (22 ms of VOT), which is a shift away from the Dutch-like boundary location and therefore unexpected.

One possible explanation for the bilinguals' category shift on higher VOT values compared to monolinguals is that the children's voicing systems are still developing. When comparing the location of the category boundaries of the children tested in Experiment II to those of the adults tested in Experiment I, the following pattern emerges: the category boundary of the adults occurs on average 12 ms before that of the monolingual children in both Dutch and German. The bilingual children's category boundary is located on even higher VOT values than those of the monolingual children in both Dutch and German. This observation indicates a developmental shift of the voicing boundary towards lower VOT values in both Dutch and German in the course of language development. A location shift of category boundaries towards a more adult-like boundary location has recently been observed in monolingual English-speaking children, but not in sequential bilingual children who were both tested approximately at age 4;4 and 5;4 (McCarthy et al., 2014). As the bilingual children in McCarthy et al. (2014) and in the present study had overall less exposure to each of their languages compared to monolinguals, they possibly take longer to acquire adult-like voicing categories in perception.

In opposition to the bilingual children who perceived VOT language-specifically, there were no detectable differences between the Dutch and German monolingual children's category boundaries or perception slopes. Given that the same monolingual children whose perception is reported here differ in their VOT production for both 'voiceless' and 'voiced' plosives (Stoehr et al., 2017b / Chapter 2), it is likely that the absence of a significant difference between the Dutch and German monolingual children's voicing perception can be ascribed to the synthetic nature of the stimulus materials used in the present experiment.

Synthesized CV VOT continua are especially useful to determine the isolated effect of VOT in voicing perception and they have successfully been used in previous research with children (Burnham et al., 1991; Hoonhorst et al., 2011; Wolf, 1973) and adults (Abramson & Lisker, 1973; Caramazza et al., 1973; Garcia-Sierra et al., 2009; Lisker & Abramson, 1970; Williams, 1977). On the downside, the lack of secondary cues makes synthetic speech more difficult to process than natural speech (Humes, Nelson & Pisoni, 1991; Logan, Greene & Pisoni, 1989; Luce, Feustel & Pisoni, 1983). It has recently been argued that children are less flexible than adults in their use of limited speech cues (Hazan & Barrett, 2000), which likely increased the task demands in the present study. It is possible, however, that bilingual children are more flexible than monolingual children in their use of limited speech cues, and were therefore better able than their monolingual peers to understand and categorize the synthetic stimuli. One reason that may lead to a bilingual advantage is that bilingual children are exposed to more variable and non-native speech on a regular basis, which is also believed to be the reason that bilingual infants and bilingual adults are more sensitive to non-native speech contrasts than age-matched monolinguals (Byers-Heinlein & Fennell, 2014; Chang, 2016; Petitto, Berens, Kovelman, Dubins, Jasinska & Shalinsky, 2012). Although accented speech and synthesized speech are qualitatively different, it is possible that exposure to more variable input is also beneficial in processing synthetic speech.

Children who had difficulty processing the synthetic speech stimuli and children for whom the task demands were too high are expected to perform poorly in categorizing the VOT endpoint stimuli correctly. Children who were more accurate in categorizing the end points of the VOT continuum relied more on VOT as a cue to voicing than children with lower accuracy identifying the VOT endpoints. The association between children's use of VOT as a cue to voicing and endpoint identification accuracy has two possible explanations. Random labeling of the VOT endpoints may indicate that children who exhibit this pattern do not yet rely on VOT as a cue to voicing. Alternatively, it suggests that not all children were able to process synthetic speech, or stay focused on the task, and therefore labelled all stimuli randomly.

Individual labeling patterns similar to those observed by Simon and Fourcin (1977) were present in our data. Adult-like categorical labeling was observed in 37% (22/59) children. Nearly half of the children exhibited scattered labeling (49%; 29/59), and 12% of the children labeled the stimuli in a more progressive or linear way (7/59). In line with Simon and Fourcin (1977), the monolingual children in this study showed a developmental trend in which labeling changed from scattered over progressive to categorical with increasing age. By contrast, no such age trend was visible in the bilingual children. Specifically, the bilingual children who labeled the stimuli categorically were on average younger than monolingual children with a categorical labeling pattern (4;11 vs. 5;4). These results suggest that younger bilingual children were better able than younger monolingual children to perform the task. This indicates that the bilingual children were more flexible in processing the synthetic stimuli or perhaps more at ease when performing the task.

5 Conclusions

Simultaneous bilingual children need to acquire the phoneme inventories of two languages in parallel. While some phonemes may be specific to one language only, others may be shared between the two languages and only differ in their phonetic implementation, as is the case for ‘voiceless’ and ‘voiced’ plosives in Dutch and German. In the past, bilingual children’s acquisition of the voicing contrast has predominantly been investigated through children’s speech productions. These studies collectively suggest phonetic differentiation for ‘voiceless’ plosives, but not ‘voiced’ plosives between the bilinguals two languages. However, bilingual children’s similar phonetic realization of ‘voiced’ plosives in their two languages does not rule out that they associate ‘voiced’ plosives with different VOT values in their two languages. Using an XAB speech sound perception task conducted in Dutch and in German, the present study showed that simultaneous bilingual children have language-specific category boundaries for the perception of voicing: they accept longer VOT values as instances of ‘voiced’ plosives in German than in Dutch. This finding in combination with the findings of previous speech production studies highlights the importance of combining speech production and speech perception methods to understand linguistic systems of bilingual children.

References

Abramson, A. S., & Lisker, L. (1973). Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics*, *1*, 1–8.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX Lexical Database (CD-ROM).* Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Bailey, P. J., & Haggard, M. P. (1980). Perception-production relations in the voicing contrast for initial stops in 3-year-olds. *Phonetica, 37*, 377–396.

Bergmann, C., Paulus, M., & Fikkert, P. (2012). Preschoolers' comprehension of pronouns and reflexives: The impact of the task. *Journal of Child Language, 39*, 777–803.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York.

Boersma, P., & Weenink, D. (2014). *Praat: Doing phonetics by computer*. Version 5.3. Retrieved from http://www.praat.org/

Brasileiro, I. (2009). *The effects of bilingualism on children's perception of speech sounds.* Utrecht: LOT.

Burnham, D. K., Earnshaw, L. J., & Clark, J. E. (1991). Development of categorical identification of native and non-native bilabial stops: Infants, children and adults. *Journal of Child Language, 18*, 231–260.

Byers-Heinlein, K., & Fennell, C. T. (2014). Perceptual narrowing in the context of increased variation: Insights from bilingual infants. *Developmental Psychology, 56*, 274–291.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America, 54*, 421–428.

Chang, C. B. (2016). Bilingual perceptual benefits of experience with a heritage language. *Bilingualism: Language and Cognition, 19*, 791–809.

Deuchar, M., & Clark, A. (1996). Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics, 24*, 351–365.

Elman, J. L., Diehl, R. L., & Buchwald, S. E. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America, 62*, 971–974.

Ferjan Ramírez, N., Ramírez, R. R., Clarke, M., Taulu, S., & Kuhl, P. K. (2016). Speech discrimination in 11-month-old bilingual and monolingual infants: A magnetoencephalography study. *Developmental Science, 20*, 1–16.

Fikkert, P. (2010). Developing representations and the emergence of phonology: Evidence from perception and production. In C. Fougeron, B. Kuehnert, M. Imperio & N. Vallee (Eds.), *Laboratory Phonology, 10* (pp. 227–260). Berlin: De Gruyter Mouton.

Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication, 6*, 185–202.

Forrest, K. & Rockman, B. K. (1988). Acoustic and perceptual analysis of word-initial stop consonants in phonologically disordered children. *Journal of Speech and Hearing Research, 31*, 449–459.

Garcia-Sierra, A., Diehl, R. L., & Champlin, C. (2009). Testing the double phonemic boundary in bilinguals. *Speech Communication, 51*, 369–378.

García-Sierra, A., Ramírez-Esparza, N., Silva-Pereyra, J., Siard, J., & Champlin, C. A. (2012). Assessing the double phonemic representation in bilingual speakers of Spanish and English: An electrophysiological study. *Brain & Language, 121*, 194–205.

Garcia-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., Ortiz, S., & Kuhl, P. K. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. *Journal of Phonetics, 39*, 546–557.

Giezen, M. R. (2011). *Speech and sign perception in deaf children with cochlear implants.* Utrecht: LOT.

Gonzales, K., & Lotto, A. J. (2013). A bafri, un pafri: Bilinguals' pseudoword identifications support language-specific phonetic systems. *Psychological Science, 24*, 2135–2142.

Haggard, M., Ambler, S., & Callow, M. (1970). Pitch as a voicing cue. *Journal of the Acoustical Society of America, 47*, 613–617.

Hazan, V. L., Barrett, S. (2000). The development of phonemic categorization in children aged 6-12. *Journal of Phonetics, 28*, 377–396.

Hazan, V. L., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech, 36*, 17–38.

Hoonhorst, I., Medina, V., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2011). Categorical perception of voicing, colors and facial expressions: A developmental study. *Speech Communication, 53*, 417–430.

Humes, L. E., Nelson, K. J., & Pisoni, D. B. (1991). Recognition of synthetic speech by hearing-impaired elderly listeners. *Journal of Speech and Hearing Research, 34*, 1180–1184.

Johnson, C. E., & Wilson, I. L. (2002). Phonetic evidence for early language differentiation: Research issues and some preliminary data. *International Journal of Bilingualism, 6*, 271–289.

Kewley-Port, D., & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics, 2*, 194–210.

Khattab, G. (2000). VOT in English and Arabic bilingual and monolingual children. In D. Nelson & P. Foulkes (Eds.), *Leeds Working Papers in Linguistics and Phonetics, 8* (pp. 95–122). Retrieved from http://www.leeds.ac.uk/arts/homepage/387/leeds_working_papers_in_linguistics_and_phonetics

Klatt, D. H. (1980). Software for a cascade/parallel synthesizer. *Journal of the Acoustical Society of America, 67*, 971–995.

Lein, T., Kupisch, T., & van de Weijer, J. (2016). Voice onset time and global foreign accent in German-French simultaneous bilinguals during adulthood. *International Journal of Bilingualism, 20*, 732 –749.

Liberman, A. M., Delattre, P. C., & Cooper, F. S. (1957). Some cues for the distinction between voiced and voiceless stops in initial position. *Language and Speech, 1*, 153–167.

Liberman, A. M., Harris, K., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*, 358–368.

Lisker, L. (1975). Is it VOT or first-formant transition detector? *Journal of the Acoustical Society of America, 57*, 1547–1551.

Lisker, L. (1978). In qualified defense of VOT. *Language and Speech, 21*, 375–383.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384–422.

Lisker, L., & Abramson, A. S. (1970). The voicing dimension: Some experiments in comparative phonetics. *Proceedings of the Sixth International Congress of Phonetic Sciences, Prague, 1967* (pp. 563–567). Prague: Academia.

Logan, J. S., Greene, B. G., & Pisoni, D. B. (1989). Segmental intelligibility of synthetic speech produced by rule. *Journal of the Acoustical Society of America, 86*, 566–581.

Luce, P. A., Feustel, T. C., & Pisoni, D. B. (1983). Capacity demands in short-term memory for synthetic and natural speech. *Human Factors, 25*, 17–32.

Macken, M. A., & Barton, D. (1980a). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language, 7*, 41–74.

Macken, M. A., & Barton, D. (1980b). The acquisition of the voicing contrast in Spanish: A phonetic and phonological study of word-initial stop consonants. *Journal of Child Language, 7,* 433–458.

MacLeod, A. A. (2016). Phonetic and phonological perspectives on the acquisition of voice onset time by French-speaking children. *Clinical Linguistics & Phonetics, 30,* 584–598.

MacLeod, A. A., & Stoel-Gammon, C. (2009). The use of voice onset time by early bilinguals to distinguish homorganic stops in Canadian English and Canadian French. *Applied Psycholinguistics, 30,* 53–77.

MacLeod, A. A., & Stoel-Gammon, C. (2010). What is the impact of age of second language acquisition on the production of consonants and vowels among childhood bilinguals? *International Journal of Bilingualism, 14,* 400–421.

Mayr, R., & Siddika, A. (2016). Inter-generational transmission in a minority language setting: Stop consonant production by Bangladeshi heritage children and adults. *International Journal of Bilingualism*. Advance online publication. doi: 10.1177/1367006916672590

McCarthy, K. M., Mahon, M., Rosen, S., & Evans, B. G. (2014). Speech perception and production by sequential bilingual children: A longitudinal study of voice onset time acquisition. *Child Development, 85,* 1965–1980.

McMurray, B. (in preparation). KlattWorks: A [somewhat] new systematic approach to formant-based speech synthesis for empirical research.

Menyuk, P., & Anderson, S. (1969). Children's identification and reproduction of /w/, /r/, and /l/. *Journal of Speech, Language, and Hearing Research, 12,* 39–52.

Nittrouer, S., & Miller, M. E. (1997). Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America, 101,* 2253–2266.

Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The "Perceptual Wedge" hypothesis as the basis for bilingual babies' phonetic processing advantage: New insights from fNIRS brain imaging. *Brain & Language, 121,* 130–143.

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

Reetz, H. (2014). *Reelex – Reetz-CELEX interface.* Version 0.4.4. Retrieved from http://web.phonetik.uni-frankfurt.de/simplex.html

Repp, B. H. (1979). Relative amplitude of aspiration noise as a voicing cue for syllable-initial stop consonants. *Language and Speech, 22*, 173–189.

Simon, C., & Fourcin, A. J. (1978). Cross-language study of speech-pattern learning. *Journal of the Acoustical Society of America, 63*, 925–935.

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017a). Second language attainment and first language attrition: The case of VOT in immersed Dutch-German late bilinguals. *Second Language Research, 33*, 483–518.

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017b). Heritage language exposure impacts voice onset time of Dutch-German simultaneous bilingual preschoolers. *Bilingualism: Language and Cognition*. Advance online publication. doi: 10.1017/S1366728917000116

Summerfield, Q. (1981). Articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 1074–1095.

Summerfield, Q., & Haggard, M. (1977). On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. *Journal of the Acoustical Society of America, 62*, 435–448.

Sundara, M., Polka, L., & Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition, 9*, 97–114.

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*, 86–110.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development, 1*, 197–234.

Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics, 37*, 35–44.

Williams, L. (1977). The perception of stop consonant voicing by Spanish-English bilinguals. *Perception & Psychophysics, 21*, 289–297.

Wolf, C. G. (1973). The perception of stop consonants by children. *Journal of Experimental Child Psychology, 16*, 318–331.

Zhou, W. (2015). Assessing birth language memory in young adoptees. Doctoral Dissertation, Radboud University Nijmegen, The Netherlands.

Zlatin, M. A., & Koenigsknecht, R. A. (1975). Development of the voicing contrast: Perception of stop consonants. *Journal of Speech and Hearing Research, 18*, 541–553.

Appendix 1. Labial and coronal plosives

Monolinguals

Model R code: Voicing Perception ~ VOT * Language * PoA + Endpoint Accuracy * VOT (1 + VOT + PoA | Child)

PoA coding: labial=1; coronal=-1

Number of observations: 1578

	β	*SE*	*z*	*p*
Intercept	-0.182	0.101	-1.809	.071
VOT	0.041	0.007	5.477	<.001
Language	-0.061	0.102	-0.597	>.250
P o A	-0.107	0.075	-1.424	.154
Endpoint Accuracy	-0.016	0.007	-2.254	.024
VOT x Language	-0.004	0.008	-0.562	>.250
VOT x P o A	0.003	0.004	0.943	>.250
Language x P o A	-0.040	0.075	-0.530	>.250
VOT x Endpoint Accuracy	0.002	0.001	3.447	<.001
VOT x Language x P o A	0.002	0.004	0.547	>.250

Bilinguals

Model R code: Voicing Perception ~ VOT * Language * PoA + Endpoint Accuracy * VOT + Exposure * Language (1 + Language + VOT + PoA | Child)

Number of observations: 1507

	β	*SE*	*z*	*p*
Intercept	-0.298	0.090	-3.308	<.001
VOT	0.039	0.007	5.335	<.001
Language	-0.065	0.064	-1.028	>.250
P o A	-0.225	0.062	-3.619	<.001
Endpoint Accuracy	-0.016	0.005	-3.056	.002
Exposure	-0.003	0.005	-0.608	>.250
VOT x Language	0.003	0.004	0.839	>.250
VOT x P o A	0.003	0.004	0.927	>.250
Language * P o A	-0.047	0.058	-0.803	>.250
VOT x Endpoint Accuracy	0.001	0.001	3.408	<.001
Language x Exposure	0.008	0.004	1.707	.088
VOT x Language x P o A	-0.003	0.004	-0.812	>.250

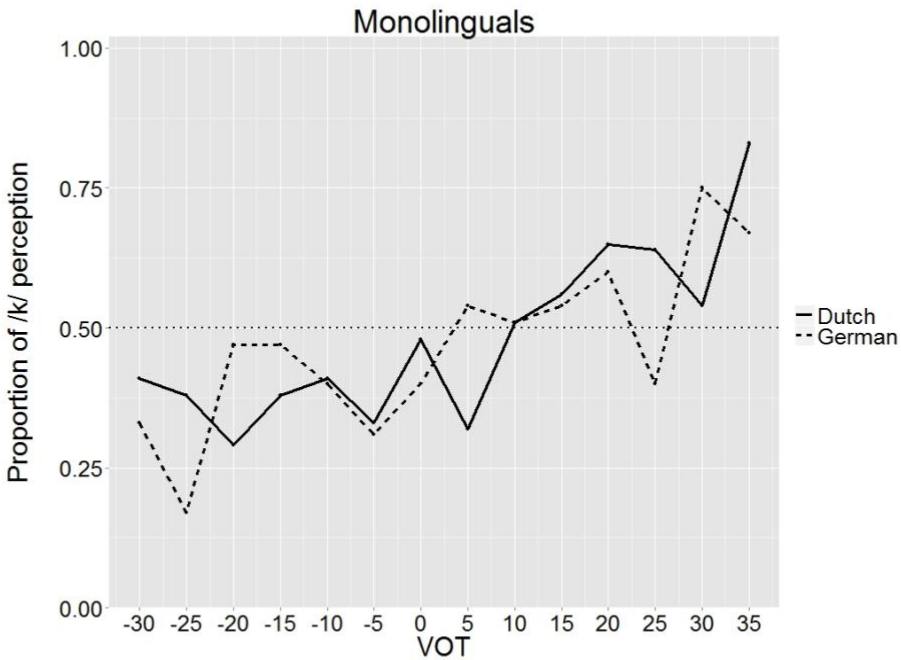
Appendix 2. Dorsal plosives

Monolinguals

Model R code: Voicing Perception ~ VOT * Language + Endpoint Accuracy * VOT (1 + VOT | Child)

Number of observations: 838

	β	*SE*	*z*	*p*
Intercept	-0.210	0.085	-2.473	.013
VOT	0.027	0.005	5.259	<.001
Language	0.012	0.086	0.139	>.250
Endpoint Accuracy	-0.009	0.006	-1.337	.181
VOT x Language	0.001	0.005	0.176	>.250
VOT x Endpoint Accuracy	0.001	0.001	3.238	.001

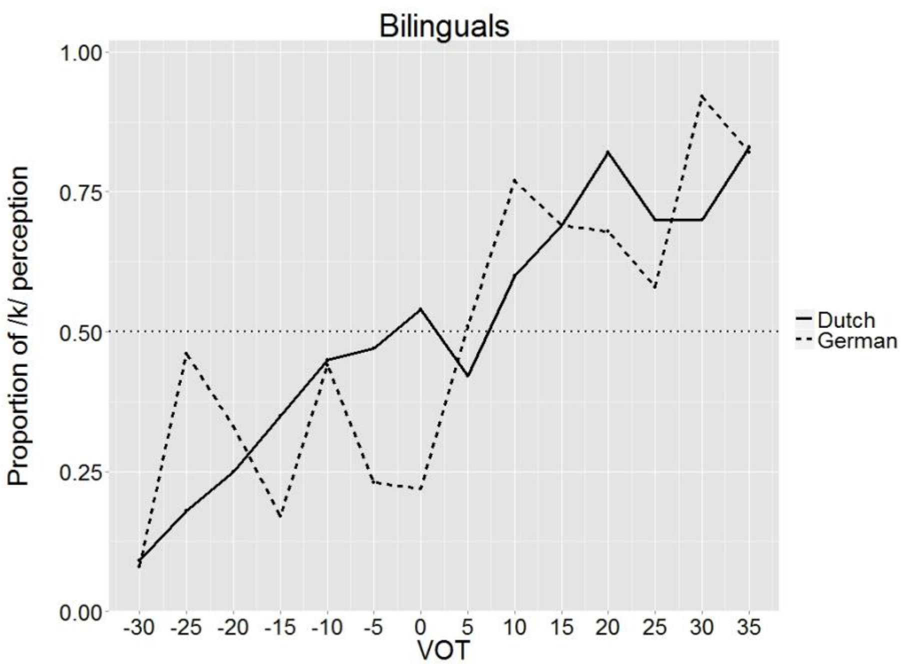


Bilinguals

Model R code: Voicing Perception ~ VOT * Language + Endpoint Accuracy * VOT (1 + Language + VOT | Child)

Number of observations: 686

	β	*SE*	*z*	*p*
Intercept	-0.176	0.098	-1.804	.071
VOT	0.053	0.010	5.484	<.001
Language	-0.104	0.097	-1.071	>.250
Exposure	-0.014	0.007	-1.926	.054
Endpoint Accuracy	-0.008	0.007	-1.131	>.250
VOT x Language	0.003	0.006	0.498	>.250
Language x Exposure	0.003	0.007	0.372	>.250
VOT x Endpoint Accuracy	0.001	0.001	1.116	>.250



Chapter 4

Feature generalization in Dutch-German bilingual and monolingual preschoolers

Based on:

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. *Feature generalization in Dutch-German bilingual and monolingual preschoolers*. Manuscript in preparation.

Abstract

Dutch and German both have a voicing contrast, but Dutch lacks the 'voiced' dorsal plosive /g/. Through this phonologically accidental gap, we addressed two primary questions: first, we investigated whether Dutch-speaking monolingual children and Dutch-German bilingual children aged 3;6-6;0 years generalized voicing to /g/. Second, we tested whether the same Dutch-German bilingual children's productions of /g/ in German are influenced by their Dutch voicing system despite the lack of /g/ in Dutch. Taken together, monolingual and bilingual children's productions of /g/ provide evidence for feature generalization: in Dutch, children either recombined their native voicing and place features to produce /g/ or resorted to producing familiar /k/. In German, the bilingual children's productions of /g/ were influenced by Dutch, suggesting either direct cross-linguistic influence from the Dutch voicing feature to German or feature generalization within the bilingual children's German phonological system.

1 Introduction

When children acquire a phonological system, they initially acquire broad contrasts, such as the contrast between labial and non-labial plosives (Jakobson, 1941). In the course of development, they acquire narrower contrasts, like the voicing contrast between 'voiced' plosives /bdg/ and 'voiceless' plosives /ptk/ (Brown & Matthews, 1997; Dresher, 2004; Jakobson, 1941; Pater, Stager & Werker, 2004; Rice & Avery, 1995). In phonological theory, contrasts can be captured through distinctive features (Chomsky & Halle, 1968). For example, the voicing contrast can be captured through the phonological feature [voice][12]. Although it has recently been argued that phonological features are limited to phonological descriptions of adult speech and may not play a role in children's language processing (Menn & Vihman, 2011), the present study provides evidence that phonological features are crucial to understanding the production patterns in children.

Phonologically accidental gaps provide a unique opportunity to investigate whether monolingual children's language processing, and thus their speech production, is based on phonological features or whole segments. A phonologically accidental gap refers to the absence of a phonemic contrast in a language, for instance due to historical changes to the phonological system (e.g., Iverson & Salmons, 2005). For example, Dutch contrasts labial and coronal plosives in voicing (i.e., 'voiced' /b/ versus 'voiceless' /p/; 'voiced' /d/ versus 'voiceless' /t/). The 'voiceless' dorsal plosive /k/, however, lacks its 'voiced' counterpart plosive /g/, which represents a phonologically accidental gap in the Dutch phoneme inventory. Non-native segments are presumably processed in the native phonological system (Silverman, 1992). In the case of Dutch, the non-native segment /g/ is composed of the native features [voice] and [dorsal]. Such a composition of native features is assumed to facilitate matching the target production (Jakobson, 1931).

Speech perception experiments provide evidence that adults' language processing is based on phonological features (Finley & Badecker, 2009). In a series of artificial grammar experiments, adult native speakers of English were exposed to morpho-phonological alternations based on back vowel harmony. The listeners

[12] There are different views on featural representations of the voicing contrast. While some researchers assume that the voicing contrast is represented through a single feature [voice] (e.g., Lombardi, 1995), others argue that the voicing contrast can best be described through monovalent features that differ depending on the phonetic implementation of the voicing contrast (e.g., Iverson & Salmons, 1995). In the latter view, prevoicing languages like Dutch have the feature [voice] and aspiration languages like German have the feature [spread glottis]. For the sake of simplicity, we use the notation [voice] to refer to 'voiced' plosives in both Dutch and German in this chapter.

generalized this alternation to new vowels, which were only presented during the experimental test phase. In a different study, it was shown that even in the early stages of language acquisition, infants appear to generalize phonetic or acoustic features (Maye, Weiss & Aslin, 2008). When trained on the Hindi voicing contrast between prevoicing and short lag voice onset time (VOT) at the dental place of articulation, English-acquiring eight-months-old infants were able to generalize this contrast to the dorsal place of articulation during the test phase (and vice versa). Taken together, these two studies suggest that the speech perception of English-native speakers (to be) is guided by generalizations of features.

Evidence for feature generalization in speech *production* comes from studies on Dutch adults' productions of the non-native segment /g/. Although the segment /g/ corresponds to a phonologically accidental gap in the Dutch plosive inventory of /bd ptk/, it occurs in a number of loan words in Dutch (Hamann & de Jonge, 2015; van Bezooijen & Gerritsen, 1994). Native speakers of Dutch in the Netherlands tend to produce word-initial /g/ in loan words predominantly as [g], but [k] or [x][13] are also common (Table 1).

Table 1. Production of /g/ by Dutch adults.

	Hamann & de Jonge (2015)	van Bezooijen & Gerritsen (1994)
% [g]	73	49
% [k]	21	21
% [x]	6	29

Productions of [g] show that Dutch speakers generalize the feature [voice] to the dorsal place of articulation. However, most Dutch adults are familiar with /g/ from English or another second language (L2; Ytsma, 2000). It is therefore inconclusive whether their productions of [g] are based on feature generalization within their native (L1) phonology or are instead rooted in their L2 phonology. Productions of /g/ as [k] suggest that Dutch native speakers map the non-native /g/ to the perceptually closest native category. Productions of /g/ as [x] are likely related to Dutch orthography, in which the orthographic representation of [x] is the letter <g>. To evaluate whether feature generalization within the L1 phonology is at play in speech production of Dutch native speakers, it is necessary to test native speakers who are not yet literate or exposed to an L2. Monolingual children fulfill these requirements.

[13] Including realizations as [χ] and [ɣ] depending on the dialect.

The question whether monolingual children's language processing is based on features or segments can also be extended to simultaneous bilingual children's language processing. It is well established that the phonologies of most bilingual children interact to some extent, a phenomenon known as cross-linguistic influence (CLI; Deuchar & Clark, 1996; Fabiano-Smith & Bunta, 2012; Johnson & Wilson, 2002; Kehoe, Lleó & Rakow, 2004; Khattab, 2000; Stoehr, Benders, van Hell & Fikkert, 2017 / Chapter 2). It remains unclear whether CLI only occurs between specific segments, such as between Dutch /b/ and German /b/ or whether a segment that is limited to one language, such as German /g/ can likewise be influenced by the Dutch voicing feature.

First support for CLI between features is provided by only one study on adult L2 learners of English, whose native Arabic dialect lacks the segment /p/ (Flege & Port, 1981). The speakers produced English /p/ as unaspirated, which can be ascribed to the unaspirated realization of 'voiceless' plosives in their L1. Because the Arabic segment inventory does not encompass /p/, this could be taken as evidence for feature level CLI. However, the unaspirated [p] is articulatory less complex than the aspirated target (Kewley-Port & Preston, 1974), such that unaspirated productions of [p] may also result from general articulatory restrictions rather than feature-level CLI alone. To reliably test whether CLI operates between features or segments, it is necessary to investigate language pairs in which CLI at the feature level would result in occurrences of the articulatory more complex production that cannot be explained by articulatory simplicity.

1.1 The present study

The present study investigates the production of /g/ in a Dutch context by Dutch-German bilingual and monolingual children, who are preliterate and have no exposure to foreign languages. /g/ is non-native to Dutch, but is composed of the Dutch native features [voice], shared with /b/ and /d/, and [dorsal], shared with /k/. Given that the children in this study do not yet receive L2 instruction, their productions of the non-native segment /g/ reflect processes in their L1 phonology.

Based on Dutch adults' production of /g/ (Hamann & de Jonge, 2015; van Bezooijen & Gerritsen, 1994), two predictions can be formulated for monolingual Dutch children's realizations of /g/: they either generalize the feature [voice] to the dorsal place of articulation, or resort to productions of the perceptually close native-segment /k/.

Bilinguals who acquire Dutch in addition to German, which encompasses /g/, may realize /g/ in the same way as monolinguals, namely through feature generalization to the dorsal place of articulation, resulting in productions of [g] or through segmental assimilation, resulting in productions of [k]. Alternatively, they may employ their German /g/ in a Dutch context. Moreover, bilinguals' productions of /g/ in *German* cannot be subject to segmental influence from Dutch and thus allow for addressing the question whether CLI in bilinguals' speech operates between features or whole segments.

The different predictions for bilinguals' productions of /g/ in Dutch and German can be teased apart based on differences in the phonetic implementation of voicing in the two languages (Jessen, 1998; Lisker & Abramson, 1964): in German, 'voiced' plosives typically have short lag VOT and 'voiceless' plosives are aspirated. Dutch has prevoiced 'voiced' plosives and short lag 'voiceless' plosives. If Dutch-speaking children generalize the feature [voice] from /b/ and /d/ and extend it to produce the non-native /g/, they should produce /g/ with prevoicing as they do for Dutch /b/ and /d/. However, if bilingual children use their German /g/ segment in a Dutch context, they are expected to produce Dutch and German /g/ alike. If, on the one hand, CLI between Dutch and German operates between whole segments, bilingual children should produce German /g/ in line with their monolingual German-speaking peers. If, on the other hand, CLI between Dutch and German operates on the feature level, bilingual children are expected to produce German /g/ with prevoicing and thus produce the articulatory and aerodynamically more complex structure that cannot be rooted in any 'default' production (Kewley-Port & Preston, 1974; Macken & Barton, 1980).

In this paper, we address the following three questions regarding productions of /g/ by monolingual Dutch children and bilingual Dutch-German children:

1) Do monolingual Dutch children generalize the feature [voice] to the dorsal place of articulation, resulting in productions of prevoiced [g] or do they produce non-native /g/ in line with the native segment /k/?

2) Do Dutch-German bilingual children follow the same feature generalization and/or segmental assimilation pattern as monolingual Dutch children, or do they produce their German /g/ segment in a Dutch context?

3) Are Dutch-German bilingual children's productions of /g/ in German influenced by their Dutch phonological system?

Based on previous findings supporting the existence of features in speech perception of infants (Maye et al., 2008) and adults (Finley & Badecker, 2009), we hypothesize that the productions of the non-native segment /g/ by monolingual Dutch children and Dutch-German bilingual children are guided by features. Moreover, we hypothesize that Dutch-German bilingual children produce prevoiced dorsal plosives in German, which may result from feature generalization across languages.

2 Method

2.1 Participants

The same children who were tested for the study reported in Chapter 2 participated in this study. One additional bilingual child had been excluded from the present analysis because she refused to name the /g/-initial target words in Dutch. The sample of bilingual children therefore only included 28 children (M_{age}=4;7, range 3;7–5;11; 13 female).

2.2 Materials and procedure

Productions of the 'voiced' and 'voiceless' dorsal plosives /g/ and /k/ were investigated. In addition, labial and coronal plosives were elicited for a different study (Stoehr et al., 2017 / Chapter 2). For each native segment (Dutch /k/; German /g/ and /k/), six target words were selected from developmental vocabulary lists. For the Dutch non-native segment /g/, one English loan word and three names[14] complying with Dutch phonotactics were used to elicit the children's productions. Each name referred to a character introduced to the children at the beginning of the testing session. The experimenter labelled each name ten times (Singh, Hui, Chan & Golinkoff, 2014; Woodward, Markman & Fitzsimmons, 1994), producing the initial /g/ with prevoicing. Table 2 lists all target words.

[14] We initially planned to limit the elicitation of /g/ in a Dutch context to four names. During a pilot with three monolingual Dutch-speaking children, we observed that children only successfully remembered three names. For this reason, we eliminated one female name and included the loan word ['gol] "goal" instead.

Testing took place in a quiet room at the children's homes. Parents gave informed consent and completed a language background questionnaire. The children named all target words in two different picture-naming tasks to enhance the number of produced tokens per child. These tasks were the same as the ones described in Chapter 2, section 2.2.2.

Table 2. German and Dutch target words.

German			Dutch		
Target word	Pronunciation	Translation	Target word	Pronunciation	Translation
Garten	['gaɐ̯t(ə)n]	garden	goal[15]	['gol]	goal
Gabel	['gab(ə)l]	fork	Gabi	['gabi]	name (f)
Gürtel	['gʏɐ̯t(ə)l]	belt	Gero	['geʀo]	name (m)
Gans	['gans]	goose	Gizmo	['gɪzmo]	name (m)
Geld	['gɛlt]	money			
Gurke	['gʊɐ̯kə]	cucumber			
Käse	['kɛ:zə]	cheese	kaas	['kas]	cheese
Katze	['katsə]	cat	kast	['kɑst]	cupboard
Kette	['kɛtə]	necklace	kikker	['kɪkər]	frog
Korb	['kɔɐ̯p]	basket	kip	['kɪp]	chicken
Kuh	['ku:]	cow	koe	['ku]	cow
Küken	['ky:k(ə)n]	chick	koning	['konɪŋ]	king

2.3 Recordings and VOT measurements

See Chapter 2, section 2.3 for a description of the recordings and VOT measurements. Across groups and plosives, 11% of the tokens were excluded from the analyses because they could not be unambiguously measured due to, for example, coarticulation, sound overlap, or whispering.

[15] Note that the combination of a loan word and novel words is a potential caveat in this study because familiar words (and phoneme sequences) are produced more accurately than novel words (and phoneme sequences; e.g., Munson, Edwards & Beckman, 2005). For this reason, we ran all analyses including and excluding the target word "goal" in Dutch. The pattern of results did not change when this target word was excluded. We therefore report only the statistical analyses including the target word "goal" in the results section.

3 Results

3.1 Descriptive statistics

Instances of prevoiced realizations of /g/ were present in all groups of children. Monolingual Dutch-speaking children prevoiced the highest percentage (27%, *SD*=28) of /g/ tokens, followed by the bilingual children who prevoiced 13% (*SD*=15) of their /g/ tokens in Dutch and also 13% (*SD*=21) in German. Monolingual German-speaking children only prevoiced 3% (*SD*=6) of their /g/ tokens. These percentages reflect the means calculated over children per group and language. The number of children who never prevoiced /g/ is lowest in the group of monolingual Dutch-speaking children, and highest in the group of monolingual German-speaking children. The number of bilingual children who never prevoiced /g/ is lower in German than in Dutch, and falls in between the two monolingual groups. The number of children who sometimes prevoiced /g/ was highest in the group of monolingual Dutch-speaking children and lowest in the group of monolingual German-speaking children. Again, the group of bilingual children fell in between the two monolingual groups, and the number of bilingual children who sometimes produced /g/ with prevoicing was higher in Dutch than in German. These results are displayed in Table 3.

Table 3. Distribution of children across groups who produced prevoicing.

	Dutch		German	
	Monolinguals	Bilinguals	Bilinguals	Monolinguals
Always prevoiced	1/30	0/29	0/28	0/29
Sometimes prevoiced	23/30	20/29	16/28	9/29
Never prevoiced	6/30	9/29	12/28	20/29

As expected based on previous research, monolingual Dutch-speaking children and bilingual children in both languages produced prevoicing less frequently for /g/ than for /b/ and /d/ (Cho & Ladefoged, 1999; Ohala & Riordan, 1979). Monolingual German-speaking children almost never produced /g/ with prevoicing. Figure 1 visualizes these patterns.

3.2 Bilingual and monolingual children

Bilingual children's production of prevoicing in /g/ was compared with their monolingual Dutch and German peers in two mixed effects logistic regression models (glmer), conducted independently for Dutch (612 observations) and German (802 observations). The main question was whether bilingual children prevoice /g/ less frequently than their monolingual peers in Dutch (*Dutch logistic model*) and more frequently than their monolingual peers in German (*German logistic model*). The dependent variable was Plosive Production (prevoiced=0; devoiced=1), and the between-subjects independent variable was Language Background (monolingual=-1, bilingual=1). As random effects, the model included intercepts for Child and Target Word, as well as by-Target Word random slopes for Language Background [model R code: Plosive Production ~ Language Background + (1 | Child) + (1 + Language Background | Target Word)].

Dutch logistic model. In Dutch, bilingual children prevoiced a lower proportion of /g/ tokens than their monolingual peers ($\beta_{Language\ Background}$=0.62, SE=0.22, z=2.81, p=.005). This finding parallels the prevoicing production pattern found in native segments, which bilingual children prevoiced less consistently than monolingual children (Kehoe et al., 2004; Khattab, 2000; Stoehr et al., 2017 / Chapter 2), suggesting that productions of /g/ follow similar mechanisms as productions of the native 'voiced' plosives /b/ and /d/.

German logistic model. In German, bilingual children prevoiced a higher proportion of /g/ tokens than their monolingual German peers ($\beta_{Language\ Background}$=-0.81, SE=0.32, z=-2.51, p=.012). This finding is in line with productions of /b/ and /d/, which bilingual children prevoiced more frequently than German-speaking monolingual children (Stoehr et al., 2017 / Chapter 2; see also Kehoe et al., 2004). Given the similarities between the bilingual children's production of [b], [d] and [g], the latter appears to be influenced by the children's Dutch phonological system despite the lack of /g/ in Dutch.

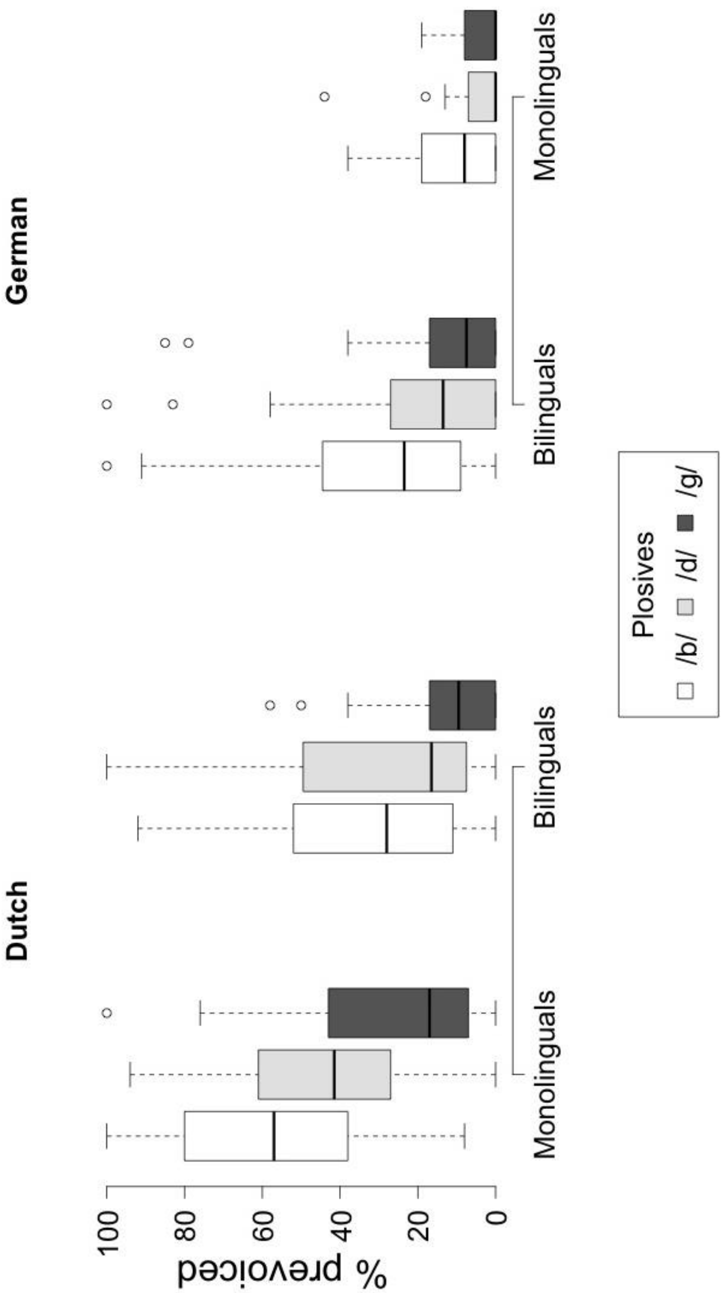


Figure 1. Percent prevoiced /g/ tokens over children by language and group. Results for /b/ and /d/ are based on Stoehr et al. (2017) / Chapter 2 of this dissertation.

The next set of analyses addressed whether Dutch-speaking and German-speaking monolingual and bilingual children produced the devoiced /g/ tokens with shorter VOT than /k/ tokens. The VOT distributions of devoiced /g/ and /k/ tokens in Figure 2 show considerable overlap in the Dutch production of bilingual and monolingual children, and little overlap in the German production of bilingual and monolingual children.

Mixed effects linear regression models (lmer) were run separately for Dutch (1234 observations) and German (1537 observations). The central question was whether the bilingual and monolingual children produced shorter VOT for the devoiced /g/ tokens compared to /k/ tokens in Dutch (*Dutch linear model*) and in German (*German linear model*). The continuous dependent variable was VOT in ms and the model included the within-subjects independent variable Voicing (/g/=-1, /k/=1) and the between-subjects independent variable Language Background (monolingual=-1, bilingual=1) with an interaction term. The model included random effects with intercepts for Child and Target Word, as well as by-Child random slopes for Voicing and by-Target Word random slopes for Language Background [model R code: VOT ~ Voicing * Language Background + (1 + Voicing | Child) + (1 + Language Background | Target Word)].

Dutch linear model. The model did not detect a difference between VOT durations of devoiced /g/ in a Dutch context compared to Dutch /k/ ($\beta_{Voicing}$=3.19, *SE*=2.89, *t*=1.10, *p*>.250; $\beta_{LanguageBackground}$=2.94, *SE*=1.53, *t*=1.91, *p*=.056; $\beta_{LanguageBackground*Voicing}$=0.38, *SE*=1.12, *t*=0.34, *p*>.250). This finding suggests that when required to produce non-native /g/ in Dutch, monolingual as well as bilingual children produce the non-prevoiced /g/ realizations in line with their Dutch /k/.

German linear model. In German, the model detected significant main effects for Voicing ($\beta_{Voicing}$=22.47, *SE*=2.66, *t*=8.45, *p*<.001) and Language Background ($\beta_{LanguageBackground}$=-3.76, *SE*=1.57, *t*=-2.40, *p*=.016), and a significant Voicing x Language Background interaction ($\beta_{Voicing*LanguageBackground}$=-5.66, *SE*=1.45, *t*=-3.90, *p*<.001). Two post hoc analyses were conducted based on the data split by group [model R code: VOT ~ Voicing + (1 + Voicing | Child) + (1 | Target Word)]. Both groups of children produced devoiced /g/ tokens with significantly shorter VOT than /k/ tokens, but the magnitude of the effect was larger in monolingual German children (monolinguals: $\beta_{Voicing}$=28.14, *SE*=2.51, *t*=11.22, *p*<.001; bilinguals: $\beta_{Voicing}$=11.88, *SE*=3.51, *t*=4.81, *p*<.001). These findings show that monolingual German children and Dutch-German bilingual children acquired a German /g/ category that is distinct from their /k/ category.

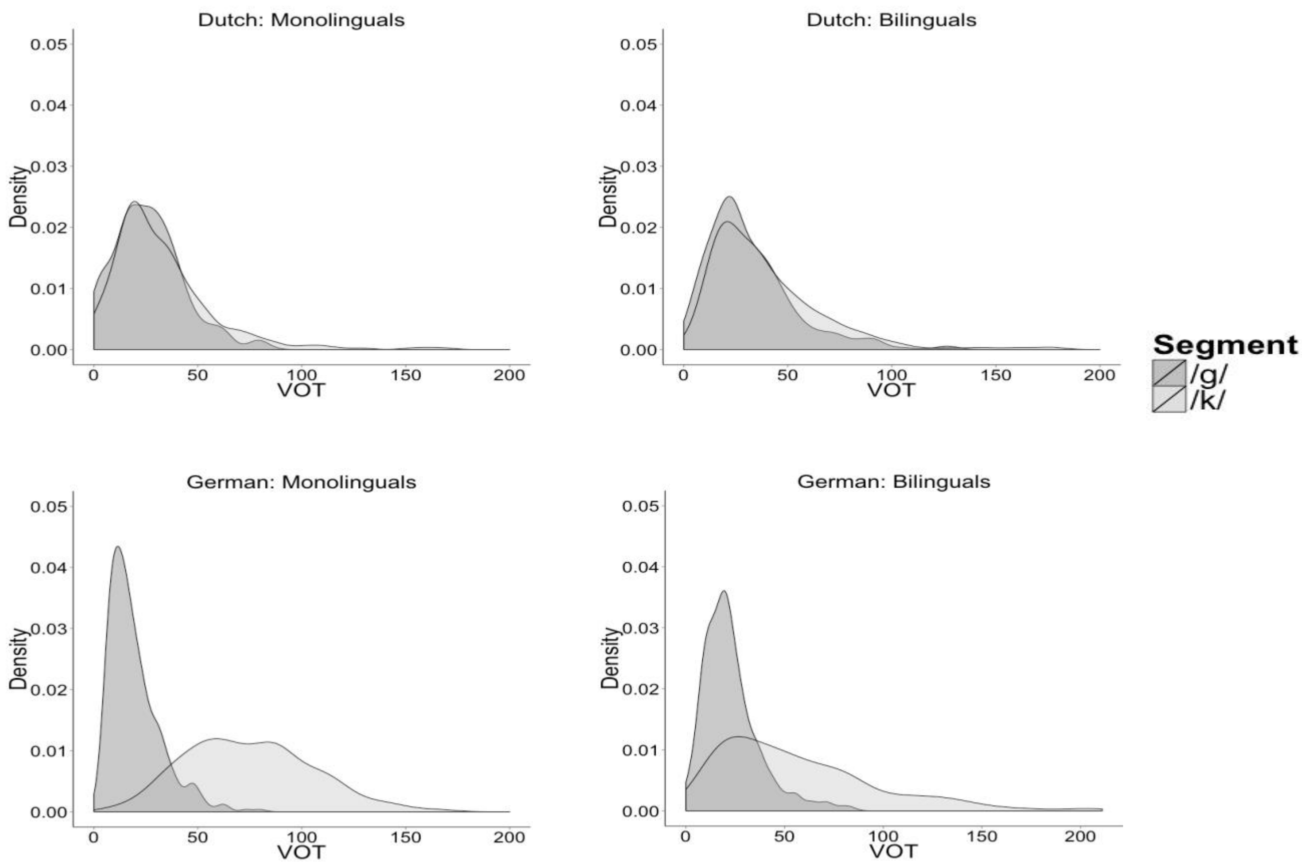


Figure 2. VOT distribution of devoiced / g/ and /k/ in Dutch (top) and German (bottom) by language background relative to each segment's density in the corpus. Results of /k/ are based on Stoehr et al. (2017) / Chapter 2 of this dissertation.

3.3 Bilingual children: Dutch and German

Our second question was whether bilingual children prevoiced a higher percentage of their /g/ tokens in Dutch than in German. A mixed effects logistic regression model (glmer; 708 observations) was run with the dependent variable Plosive Production (prevoiced=0; devoiced=1), and the within-subjects independent variable Language (Dutch=-1, German=1). The model included random effects for Child and Target Word, and by-Child random slopes for Language [model R code: Plosive Production ~ Language + (1 + Language | Child) + (1 | Target Word)]. The model did not detect a difference in the proportion of prevoiced /g/ tokens between Dutch and German ($\beta_{Language}$=0.39, *SE*=0.25, *z*=1.54, *p*=.123), raising the possibility that the bilingual children's prevoiced productions of /g/ in a Dutch context resulted from borrowing the German /g/ segment.

The majority of the bilingual children's /g/ tokens (87% in both languages) was devoiced, and a mixed effects linear regression analysis (lmer; 612 observations) tested whether the bilingual children's VOT of devoiced /g/ differed between Dutch and German. The continuous dependent variable was VOT in ms, and the within-subjects independent variable was Language (Dutch=-1, German=1). The model included random effects with intercepts for Child and Target Word, and by-Child random slopes for Language [model R code: VOT ~ Language + (1+ Language | Child) + (1 | Target Word)]. This analysis revealed that bilingual children produced *longer* VOT for devoiced /g/ in a Dutch context than for German devoiced /g/ ($\beta_{Language}$=-4.28, *SE*=1.68, *t*=-2.56, *p*=0.12). This finding suggests that the bilingual children's production of /g/ in a Dutch context cannot be explained entirely by borrowing this segment from German. Instead, the bilingual children's production of longer VOT in Dutch /g/ than German /g/ supports the hypothesis that their devoiced Dutch /g/ production is in line with their production of Dutch /k/.

4 Discussion

This study investigated productions of /g/ in a Dutch context by monolingual Dutch children and Dutch-German bilingual children to address two overarching questions: first, do monolingual Dutch children generalize the feature [voice] to the dorsal place of articulation, resulting in productions of prevoiced [g] or do they produce non-native /g/ in line with their native segment /k/? Second, do Dutch-German bilingual children follow production patterns similar to monolingual Dutch children or do they produce their German /g/ segment in a Dutch context? Moreover, as a third question, the bilingual children's productions of /g/ in German were investigated to address the questions whether their productions of /g/ in German are influenced by their Dutch phonological system despite the lack of /g/ in Dutch. Investigations of all three questions point toward the existence of features in children's language processing, as discussed below.

4.1 Production of the non-native segment /g/ in Dutch

Productions of /g/ in a Dutch context can reveal whether the language processing of Dutch-speaking children is mediated by phonological features. Monolingual Dutch-speaking children prevoiced approximately one quarter of their /g/ tokens, and this proportion is lower than their proportion of prevoiced /b/ and /d/ tokens (see Stoehr et al., 2017 / Chapter 2). This production pattern is in line with the expected decrease in the production of prevoicing the further the plosive's place of articulation is to the back of the mouth, as observed in adults across a variety of languages (Cho & Ladefoged, 1999; Ohala & Riordan, 1979).

The monolingual children's production of prevoiced [g] suggests generalization of the feature [voice], which is phonetically implemented with prevoicing in Dutch, to the dorsal place of articulation. Feature generalization had previously been observed in speech perception of English-acquiring infants (Maye et al., 2008) and English-speaking adults (Finley & Badecker, 2009). Moreover, productions of [g] in loan words by Dutch-speaking adults had been reported previously (Hamann & De Jonge, 2015; van Bezooijen & Gerritsen, 1994). Given that virtually all Dutch adults have some L2 knowledge (Ytsma, 2000), their production of [g] may have been influenced by L2 phonologies. According to parental report, the monolingual Dutch children in this study were not yet exposed to an L2, suggesting that generalization of the voicing contrast to the dorsal place of articulation can account for their productions.

The bilingual children prevoiced fewer /g/ tokens (13%) in a Dutch context than their monolingual Dutch peers (27%). This finding is expected given that the bilingual children also prevoiced the native Dutch segments /b/ and /d/ less consistently than their monolingual Dutch-speaking peers (Stoehr et al., 2017 / Chapter 2). There are two possible explanations for the bilingual children's production of prevoiced [g] in a Dutch context. First, borrowing /g/ from German, in which the bilingual children also prevoiced 13% of all /g/ tokens, can account for their productions of prevoiced [g] in a Dutch context. Second, the bilingual children's productions of prevoiced [g] in a Dutch context can be explained by the same feature generalization processes that appears to mediate monolingual Dutch-speaking children's production of prevoiced [g]. To follow up on these two possible explanations, we investigated VOT durations of *devoiced* productions of [g] by bilingual and monolingual children.

Monolingual and bilingual Dutch-speaking children predominantly produced the /g/-initial target words without prevoicing (73% and 87% of all tokens, respectively). The VOT durations of these tokens were indistinguishable from the children's VOT durations of the 'voiceless' dorsal plosive /k/ (see Stoehr et al., 2017 / Chapter 2). These similarities in VOT duration suggest that monolingual and bilingual Dutch-speaking children resorted to a known feature combination matching the native Dutch segment /k/. It is likely that production of /g/ resembling /k/ in VOT is rooted in children's assimilation of non-native /g/ with their native Dutch /k/ category (Silverman, 1992).

When comparing bilingual children's VOT durations in devoiced /g/ tokens produced in a Dutch context to their productions of VOT in devoiced /g/ tokens in German, it was observed that the bilingual children produced significantly *longer* VOT in Dutch than in German. If the bilingual children had aimed to produce their German /g/ in a Dutch context, but failed to do so with prevoicing for aerodynamic reasons (Cho & Ladefoged, 1999; Ohala & Riordan, 1979), they should have produced similar VOT durations in Dutch and in German. As the results were to the contrary, these findings suggest that the bilingual children's productions of prevoiced and devoiced /g/ in a Dutch context can best be explained by feature generalization within their Dutch phonological system.

In sum, productions of /g/ in a Dutch context by monolingual and bilingual children provide evidence for the existence of features in their still-developing phonology. Both groups of children follow the same patterns and either combine the two native features [voice] and [dorsal] into a novel segment /g/ or assimilate /g/ to the 'voiceless' native segment /k/, which shares the place feature [dorsal] with /g/. Most children apply both strategies, suggesting a lack of automation due to little experience with /g/. General uncertainty in categorization of non-native segments is also common in adults' speech perception (Tyler, Best, Faber & Levitt, 2014).

4.2 Production of the native segment /g/ in German

Dutch-German bilingual children's production of German /g/ compared to monolingual German children can provide insight into the nature of CLI. Previous research reported CLI in bilingual children acquiring languages that differ in the phonetic implementation of voicing, as is the case for Dutch and German (Deuchar

& Clark, 1996; Fabiano-Smith & Bunta, 2012; Johnson & Wilson; Kehoe et al., 2004; Khattab, 2000). These studies included only productions of segments that are native to both languages of the bilingual children and did therefore not address whether CLI occurs directly between specific segments or between features.

The present study showed that Dutch-German bilingual children prevoiced more /g/ tokens than their monolingual German-speaking peers, who virtually never prevoiced /g/. This consistently higher proportion of prevoiced /g/ tokens in the bilingual children's German cannot be explained by CLI at the segment-level as the Dutch phoneme inventory lacks /g/. Rather, the phonological feature [voice] of Dutch /b/ and /d/, which is phonetically implemented with prevoicing, seems to influence the bilingual children's production of German /g/ in the same way as it influences their German /b/ and /d/ (see Stoehr et al., 2017 / Chapter 2).

In the present data, CLI from Dutch to German caused the bilingual children to produce the articulatory and aerodynamically *more* complex prevoicing, which is typically acquired late by monolingual children (Kewley-Port & Preston, 1974; Macken & Barton, 1980). The bilingual children's production of prevoicing in German /g/ is especially remarkable because prevoicing is difficult to produce at the dorsal place of articulation (Cho & Ladefoged, 1999; Ohala & Riordan, 1979). As the more complex rather than the simpler structure was found in the bilingual children's German production, articulatory restrictions can be ruled out as a competing hypothesis to CLI. These findings corroborate previous findings on adult L2-English speakers whose native Arabic dialect lacks /p/, and who produced English /ptk/ with a lack of aspiration, which can be ascribed to their L1 phonology (Flege & Port, 1981).

The influence of the Dutch voicing feature on the bilingual children's production of [g], however, can also be explained by feature generalization within the children's German phonological system. It is evident that the bilingual children's productions of German [b] and [d] are influenced by their Dutch phonological system, be it at the segment or feature level (see Stoehr et al. 2017 / Chapter 2). Similar to the monolingual and bilingual children's feature generalization in Dutch, the bilingual children's production of [g] in German may likewise result from feature generalization from their prevoiced German /b/ and /d/ to / g/.

Although the bilingual children produced some /g/ tokens with prevoicing in German, they produced most /g/ tokens in German without prevoicing, as did their monolingual German-speaking peers. These devoiced /g/ tokens, which bilingual and monolingual children produced in German, had significantly shorter VOT than their /k/ tokens. This VOT difference between /g/ and /k/ shows that monolingual

German and bilingual children maintained a clear contrast between devoiced /g/ and /k/ when speaking German. In sum, the bilingual children acquired a /g/ category in German, which encompasses optional prevoicing and short lag VOT. This /g/ category is distinct from their aspirated /k/ category, which shows that the absence of a voicing contrast in Dutch did not hinder their acquisition of the voicing contrast at the dorsal place of articulation in German.

5 Conclusions

This study provided evidence for the existence of phonological features within preliterate children's native phonology. Monolingual Dutch children generalized the feature [voice] to the dorsal place of articulation, suggesting that features play a role in children's language processing. However, this generalization is not consistent, and most of the time, Dutch-speaking children resort to productions of the perceptually and featurally close native-segment /k/. The inconsistent use of both strategies may result from a lack of automation in the production of non-native segments. Dutch-German bilingual children appear to be following the same production patterns of /g/ in a Dutch context as their monolingual Dutch-speaking peers, although they are familiar with /g/ from German. Moreover, Dutch-German bilingual children's production of German /g/ reflects phonological influence from their majority language Dutch, suggesting that CLI may affect segments that only occur in one of the bilinguals' languages. In this specific case, the bilingual children's production of prevoiced [g] in German may be based on feature level CLI from Dutch to German or on feature generalization within the children's German phonological system. The latter would indicate an indirect influence of the Dutch phonological system. It is important to stress that the bilingual children's productions of /g/ in both Dutch and German can be explained by feature generalization processes within their Dutch or German phonological system, respectively. This suggests that bilingual children may function in a monolingual-like fashion.

References

Brown, C., & Matthews, J. (1997). The role of feature geometry in the development of phonological contrasts. In S. J. Hannahs & M. Young-Scholten (Eds.), *Focus on phonological acquisition* (pp. 67–112). Amsterdam: Benjamins.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics, 27*, 207–229.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York, NY: Harper & Row.

Deuchar, M., & Clark, A. (1996). Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics, 24*, 351–365.

Dresher, B. E. (2004). On the acquisition of phonological contrasts. In J. van Kampen & S. Baauw (Eds.), *Proceedings of Generative Approaches to Language Acquisition 2003. Volume 1* (pp. 27–46). Utrecht: LOT.

Fabiano-Smith, L., & Bunta, F. (2012). Voice onset time of voiceless bilabial and velar stops in 3-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics & Phonetics, 26*, 148–163.

Finley, S., & Badecker, W. (2009). Artificial language learning and feature-based generalization. *Journal of Memory and Language, 61*, 423–437.

Flege, J. E., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech, 24*, 125–146.

Hamann, S., & de Jonge, A. (2015). Eliciting the Dutch loan phoneme /g/ with the menu task. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow: University of Glasgow.

Iverson, G. K., & Salmons, J. C. (1995). Aspiration and laryngeal representation in Germanic. *Phonology, 12*, 369–396.

Iverson, G. K., & Salmons, J. C. (2005). Filling the gap: English tense vowel plus final /š/. *Journal of English Linguistics, 33*, 207–221.

Jakobson, R. (1931). Prinzipien der historischen Phonologie. *Traveaux du Cercle Linguistique de Prague, 4*, 247–267.

Jakobson, R. (1941). *Kindersprache, Aphasie und allgemeine Lautgesetze.* Uppsala: Almqvist & Wiksell.

Jessen, M. (1998). *Phonetics and phonology of tense and lax obstruents in German.* Amsterdam: Benjamins.

Johnson, C., & Wilson, I. (2002). Phonetic evidence for early language differentiation: Research issues and some preliminary data. *International Journal of Bilingualism, 6*, 271–289.

Kehoe, M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition, 7*, 71–88.

Kewley-Port, D., & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics, 2*, 194–210.

Khattab, G. (2000). VOT in English and Arabic bilingual and monolingual children. In D. Nelson & P. Foulkes (Eds.), *Leeds Working Papers in Linguistics and Phonetics, 8* (pp. 95–122). Retrieved from http://www.leeds.ac.uk/arts/homepage/387/leeds_working_papers_in_linguistics_and_phonetics

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384–422.

Lombardi, L. (1995). Laryngeal neutralization and syllable wellformedness. *Natural Language and Linguistic Theory, 13*, 39–74.

Macken, M. A., & Barton, D. (1980). The acquisition of the voicing contrast in Spanish: A phonetic and phonological study of word-initial stop consonants. *Journal of Child Language, 7*, 433–458.

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science, 11*, 122–134.

Menn, L., & Vihman, M. (2011). Features in child phonology. Inherent, emergent, or artefacts of analysis? In G. N. Clements & R. Ridouane (Eds.), *Where do phonological features come from?: Cognitive, physical and developmental bases of distinctive speech categories* (pp. 261–301). Amsterdam: Benjamins.

Munson, B., Edwards, J., & Beckman, M. E. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language, and Hearing Research, 48*, 61–78.

Ohala, J. J., & Riordan, C. J. (1979). Passive vocal tract enlargement during voiced stops. In J. J. Wolf & D. H. Klatt (Eds.), *Speech communication papers* (pp. 89–92). New York, NY: Acoustical Society of America.

Pater, J., Stager, C., & Werker, J. (2004). The perceptual acquisition of phonological contrasts. *Language, 80*, 384–402.

Rice, K., & Avery, P. (1995). Variability in a deterministic model of language acquisition: A theory of segmental elaboration. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 23–42). Hillsdale, NJ: Lawrence Erlbaum.

Silverman, D. (1992). Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology, 9*, 289–328.

Singh, L., Hui, T. J., Chan, C., & Golinkoff, R. M. (2014). Influences of vowel and tone variation on emergent word knowledge: A cross-linguistic investigation. *Developmental Science, 17*, 94–109.

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017). Heritage language exposure impacts voice onset time of Dutch-German simultaneous bilingual preschoolers. *Bilingualism: Language and Cognition*. Advance online publication. doi: 10.1017/S1366728917000116

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica, 71*, 4–21.

van Bezooijen, R., & Gerritsen, M. (1994). De uitspraak van uitheemse woorden in het Standaard-Nederlands: Een verkennende studie. *De Nieuwe Taalgids, 87*, 145–160.

Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Science, 30*, 553–566.

Ytsma, J. (2000). Trilingual primary education in Friesland. In J. Cenoz & U. Jessner (Eds.), *English in Europe: The acquisition of a third language* (pp. 222–235). Clevedon: Multilingual Matters.

Chapter 5

Second language attainment and first language attrition: The case of VOT in immersed Dutch-German late bilinguals

Based on:

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017). Second language attainment and first language attrition: The case of VOT in immersed Dutch-German late bilinguals. *Second Language Research*, *33*, 483–518.

5

Abstract

Speech of late bilinguals has frequently been described in terms of cross-linguistic influence (CLI) from the native language (L1) to the second language (L2), but CLI from the L2 to the L1 has received relatively little attention. This article addresses L2 attainment and L1 attrition in voicing systems through measures of voice onset time (VOT) in two groups of Dutch–German late bilinguals in the Netherlands. One group comprises native speakers of Dutch and the other group comprises native speakers of German, and the two groups further differ in their degree of L2 immersion. The L1-German–L2-Dutch bilinguals (*N*=23) are exposed to their L2 at home and outside the home, and the L1-Dutch–L2-German bilinguals (*N*=18) are only exposed to their L2 at home. We tested L2 attainment by comparing the bilinguals' L2 to the other bilinguals' L1, and L1 attrition by comparing the bilinguals' L1 to Dutch monolinguals (*N*=29) and German monolinguals (*N*=27). Our findings indicate that complete L2 immersion may be advantageous in L2 acquisition, but at the same time it may cause L1 phonetic attrition. We discuss how the results match the predictions made by Flege's Speech Learning Model and explore how far bilinguals' success in acquiring L2 VOT and maintaining L1 VOT

depends on the immersion context, articulatory constraints and the risk of sounding foreign accented.

1 Introduction

Adults speaking a second language (L2) are likely to be identified as non-native speakers due to properties of their first language (L1) in their L2 speech (Brennan, Ryan & Dawson, 1975; Ferguson & Garnica, 1975; Flege, 1980, 1981; Scovel, 1969). Immersion in an L2 environment may cause the L2 to play a dominant role in everyday life, and may reduce the use of the L1 and contact to other native speakers. While L2 immersion can be beneficial to approach a native accent in the L2, the associated reduced L1 use may cause linguistic abilities in the L1 to deteriorate, a phenomenon known as L1 attrition (Freed, 1982; Schmid, 2004). When L1 attrition affects the domains of phonology or phonetics, it can surface as foreign-accented L1 speech (Bergmann, Nota, Sprenger & Schmid, 2016; de Leeuw, Schmid & Mennen, 2010; Hopp & Schmid, 2013). The present study combines investigations of L2 attainment and L1 attrition in the speech of two groups of late bilinguals who differ in their degree of L2 immersion to assess potential bidirectional L1–L2 influences in their phonetic systems.

Bidirectional L1–L2 influences in a bilingual's speech can be explained by the Speech Learning Model (SLM; Flege, 1995). The SLM postulates that bilinguals have a common L1–L2 phonetic space and that these phonetic systems remain to some degree flexible in adulthood. If an L2 sound is not perceived as sufficiently different from an L1 sound, it may be classified as this phonetically similar L1 sound, a process known as equivalence classification. As a result of equivalence classification in perception, also the speaker's production of that L2 sound may be different from native speakers' productions.

5

New L2 categories can be established provided they are perceived as sufficiently different from existing L1 sounds. Nevertheless, new L2 categories in a bilingual's L1–L2 phonetic space may still deviate from those of monolingual native speakers, for example to maintain contrasts with the bilingual's L1 categories. Hence, the speech of an L2 speaker who acquired new L2 categories may still deviate from native speech.

The SLM's assumption that phonetic systems remain flexible over the lifespan also implies that L1 categories can change under the influence of L2 acquisition, which can lead to a foreign accent in the L1. For this reason, the SLM has previously been used to interpret phonetic L1 attrition (Bergmann et al., 2016; Chang, 2012; Mayr, Price & Mennen, 2012). In order to understand how phonetic categories are organized in a speaker who accommodates two languages, it is important to characterize phonetic properties in both L2 and L1 speech (Chang, 2012; de Leeuw, Mennen & Scobbie, 2012, 2013; Flege & Eefting, 1987a, 1987b; Mayr et al., 2012; Mennen, 2004; Sancier & Fowler, 1997).

Bilinguals' linguistic skills in the L2 are typically established by comparing their speech against monolingual native speech (Abrahamsson & Hyltenstam, 2009; Bongaerts, van Summeren, Planken & Schils, 1997). If the goal is to determine to what extent bilinguals have been able to adapt to the phonetic environment in which they actually acquire the L2, a comparison against monolingual native speakers may be unsuitable (for similar thoughts on heritage language acquisition, see Rothman, 2007). For example, consider an L2 learner who acquires the L2 in the home country where he or she is exposed to other non-native speakers (e.g., non-native instructors or fellow L2 speakers in the home country) or to a native speaker with attrited L1 speech (e.g., an immigrant from the L2 country). In this case, comparing L2 speakers with monolingual native speakers implies that L2 speakers are evaluated against a type of speech to which they are barely exposed.

The monolingual reference point is also problematic because bilinguals are affected by cross-linguistic competition between their two languages (Cook, 2007; Hopp & Schmid, 2013; Kroll, Bobb & Wodniecka, 2006; Kupisch, Akpınar & Stöhr, 2013; Rothman & Treffers-Daller, 2014; Schmid, Gilbers & Nota, 2014). In addition, bilinguals presumably have to accommodate more phonetic categories than monolinguals. For example, consider a native speaker of Dutch who acquired German as L2 and a monolingual native speaker of German. The L2 speaker's phonetic system comprises L1-Dutch and presumably L2-German sounds, while the monolingual's phonetic system only comprises L1-German sounds. The mere process of becoming bilingual, with more phonetic categories to accommodate, may make the monolingual state impossible to attain. If we aim to test to what extent L2 speakers approach the speech of their linguistic environment, both the characteristics of the language to which they are exposed and the fact that they are bilingual need to be acknowledged. These two considerations make it important to compare bilinguals to native speakers who have been exposed to a comparable linguistic environment and who are bilinguals themselves (Cook, 2007; Hopp & Schmid, 2013; Kroll et al., 2006; Kupisch et al., 2013; Rothman & Treffers-Daller, 2014; Schmid et al., 2014).

A bilingual's daily linguistic environment is largely determined by the country of residence and may influence the linguistic skills in both L1 and L2. Bilinguals immersed in the L2 country are likely to be exposed to more speakers of their L2 compared to L2 speakers who live in their home country. The number of speakers who provide linguistic input has recently been identified as an important factor in the early stages of monolinguals' phonotactic learning (Seidl, Onishi & Cristia, 2014) and heritage speakers' lexical development (Gollan, Starr & Ferreira, 2015). Furthermore, quality and quantity of native language input play a crucial role in maintaining a native-like L1 accent after immigration to an L2 country (de Leeuw et al., 2010; Mayr et al., 2012). Input quality, quantity and diversity as captured through the country of residence are possibly also crucial factors in L2 acquisition.

The present study specifically focuses on the production of voice onset time (VOT) in two groups of late bilingual adults who live in binational households either in their home country or the L2 country, and who are L2 speakers and potentially L1 attriters. VOT is an acoustic cue that can contribute to a perceived foreign accent in both L2 speakers and L1 attriters (Flege, 1984; Flege & Eefting, 1987b; Major, 1987; Riney & Takagi, 1999; Sancier & Fowler, 1997; Schoonmaker-Gates, 2015). The present research enriches the existing literature on VOT in L2 attainment and L1 attrition in three important ways. First, it implements the methodological considerations on L2 attainment outlined above by evaluating L2 speech against the speech of native speakers who are bilinguals themselves and whose speech is characteristic to the L2 speakers' linguistic environment. Second, it brings together investigations of L2 attainment and L1 attrition in the same speakers. Third, the present experiments cover VOT production in 'voiceless' and 'voiced' plosives to allow insight into the speakers' voicing contrasts. By addressing these three considerations, the present study allows assessing the possible restructuring of bilinguals' voicing systems.

VOT is the most important acoustic cue to distinguish 'voiced' and 'voiceless' plosives, and describes the time interval between a plosive's burst release and the onset of voicing (Abramson & Lisker, 1973; Lisker & Abramson, 1964). The VOT continuum can be divided into three phonetic categories: prevoicing (negative VOT), short lag (short positive VOT) and aspiration (long positive VOT). Dutch contrasts prevoiced 'voiced' and short lag 'voiceless' plosives (e.g., Lisker & Abramson, 1964). German contrasts short lag 'voiced' and aspirated 'voiceless' plosives (e.g., Jessen, 1998). Thus, depending on the language, short lag plosives can be phonologically classified as 'voiceless' (in Dutch) or 'voiced' (in German). Although 'voiced' plosives do not require prevoicing in German, adult native speakers sometimes prevoice initial singleton plosives (Fischer-Jørgensen, 1976; Hamann & Seinhorst, 2016; Jessen, 1998; Kohler, 1977; Stock, 1971).

In production, prevoicing, short lag and aspiration differ in the required velopharyngeal activity, which is reflected in children's acquisition order (Allen, 1985; Bortolini, Zmarich, Fior & Bonifacio, 1995; Kager, van der Feest, Fikkert, Kerkhoff & Zamuner, 2007; Kewley-Port & Preston, 1974; Khattab, 2000; Macken & Barton, 1980a, 1980b; MacLeod, 2016; Stoehr, Benders, van Hell & Fikkert, 2017 / Chapter 2): across different languages, children produce the least complex short lag VOT in their early babbles. Around their second birthday, children acquiring an aspiration language produce aspiration, for which the glottis must remain open throughout consonantal closure. Substantially later, possibly in the early school years, children speaking a prevoicing language attain adult-like prevoicing, for which the glottis must be closed considerably before consonantal release and, additionally, vocal fold vibration must be initiated and sustained (Kewley-Port & Preston, 1974).

Within each phonetic category, small VOT differences can arise depending on the consonantal place of articulation (e.g., Lisker & Abramson, 1964) and, in the case of 'voiceless' aspirated plosives, word length (Flege, Frieda, Walley & Randazza, 1998; Yu, De Nil & Pang, 2015). In addition, male speakers produce optional prevoicing more frequently than female speakers (Ryalls, Zipprer & Baldauff, 1997), which can be ascribed to gender-related differences in vocal tract morphology (Fitch & Giedd, 1999).

1.1 Previous research into VOT in L2 acquisition

When bilinguals speak two languages that implement the voicing contrast differently, as is the case for the participants in the present study, a potential influence from L1 to L2 can be measured in their VOT. For 'voiceless' plosives, three different acquisition patterns have been observed in late bilinguals whose L1 is a prevoicing language (Arabic, Dutch, French or Spanish) and who learn an aspiration L2 (English or German): (1) native-like acquisition (Schmid et al., 2014; Simon, 2009; Simon & Leuschner, 2010, the phonetically trained participants); (2) differential acquisition (Flege, 1987, 1991; Flege & Eefting, 1987a, 1987b; Simon & Leuschner, 2010, the phonetically untrained participants); and (3) complete L1-to-L2 transfer (Flege, 1987, the least experienced participants; Flege & Port, 1981).

The native-like VOT acquisition pattern has been observed in highly advanced L1-immersed native speakers of Belgian Dutch with L2-English (and some participants with L3-German). The late bilinguals produced VOT in English (and German) 'voiceless' plosives similar to monolingual native speakers (Simon, 2009; Simon & Leuschner, 2010). Similarly, native speakers of Dutch in the Netherlands reached comparable VOT durations in English as English native speakers who were also immersed in a Dutch environment (Schmid et al., 2014).

These studies demonstrate that native-like aspiration of ‘voiceless’ plosives can be acquired without L2 immersion.

The differential VOT acquisition pattern occurs when bilinguals produce VOT differently in their L2 than in their L1, but still deviate from native speakers’ VOT in the L2. This pattern has been observed in bilinguals with L1-Spanish who learned L2-English as adults: their VOT was longer in English than in Spanish, but their English VOT was nevertheless shorter than that of monolingual English speakers (Flege, 1991). The same pattern emerged in bilinguals with L1-Spanish who learned L2-English during childhood, and occurred irrespective of whether they were immersed in an English environment or not (Flege & Eefting, 1987a). Similar results come from Dutch native speakers in the Netherlands with L2-English and L3-German who were not formally instructed in L2 and L3 phonetics. The speakers produced distinct VOT values for Dutch short lag ‘voiceless’ plosives versus English and German aspirated ‘voiceless’ plosives. Yet, their aspirated VOT productions in English and German still appeared shorter than the VOT of English and German monolinguals, although no direct statistical comparison was administered (Simon & Leuschner, 2010). L2 speakers with some level of L2 proficiency can thus differentiate L1 and L2 plosives in VOT, but do not necessarily reach native-like VOT.

The complete L1-to-L2 VOT transfer pattern has been observed in L1-Arabic speakers with L2-English in the USA (Flege & Port, 1981). Their VOT for English ‘voiceless’ plosives was similar to Arabic and was therefore shorter than the VOT of English monolinguals. Although the L2 speakers were immersed in the L2 country for several years, they did not show evidence for phonetic differentiation between L1 and L2 VOT. L2 immersion thus does not always lead to the acquisition of new – be it native-like or differential – L2 VOT for ‘voiceless’ plosives.

In sum, most studies on L2 VOT dealt with the acquisition of ‘voiceless’ plosives. For aspirated ‘voiceless’ plosives, native-like acquisition, differential acquisition, and complete L1-to-L2 transfer have been observed, as was described above. For the acquisition of short lag ‘voiceless’ plosives, native-like acquisition has never been reported, but L2 acquisition of short lag ‘voiceless’ plosives has only been addressed in one study (Flege, 1987).

Studies on late bilinguals’ production of ‘voiced’ plosives reveal two acquisition patterns: native-like acquisition and L1-to-L2 transfer. The native-like acquisition pattern has been observed for L2 short lag ‘voiced’ plosives in only one sample of Dutch native speakers with L2-English even though they were not immersed in the L2-speaking country (Schmid et al., 2014). The L1-to-L2 transfer pattern of L1 prevoicing to L2 short lag has also been observed, even in advanced and phonetically trained L2 speakers (Simon, 2009; Simon & Leuschner, 2010). Similarly, bilinguals who acquired their L2 during childhood tend to produce ‘voiced’ plosives with prevoicing in both languages, especially when their dominant language requires prevoicing (Flege & Eefting, 1987a; Hazan & Boulakia, 1993; MacLeod & Stoel-Gammon, 2009; Sundara, Polka & Baum, 2006).

No data are yet available on the opposite scenario: late bilinguals’ acquisition of L2 prevoiced ‘voiced’ plosives when their L1 does not require prevoicing. The present study fills this gap in the literature by contributing data on the production of ‘voiced’ plosives in Dutch by native speakers of German.

In sum, native-like attainment and VOT differentiation between L1 and L2 do not seem to require immersion, and do not automatically result from immersion either. Two studies suggest that VOT differentiation may instead be related to language experience. This relationship was observed for the acquisition of ‘voiceless’ plosives in bilinguals whose L1 was a prevoicing language (Spanish) learning an aspiration L2 (English), as well as in bilinguals with an aspiration L1 (English) learning a prevoicing L2 (French) (Flege, 1987; Flege & Eefting, 1987a). The more advanced L2 speakers in these two studies produced different VOT in their L2 than in their L1, but still showed differential VOT acquisition. Only the less experienced L2 speakers were affected by full L1-to-L2 transfer and thus did not produce language-specific VOT. These studies suggest that language experience contributes to differentiating VOT between L2 and L1, but it may not necessarily be a sufficient predictor for native-like VOT acquisition in the L2.

1.2 Previous research into VOT in phonetic L1 attrition

In some L2 speakers, the reverse of L1-to-L2 influence can be observed, namely an influence from L2 to L1. Bilinguals whose L2 has become the dominant language, for example through L2 immersion, are generally more prone to L1 attrition than L1-dominant bilinguals (Schmid & Köpke, 2007). The present study also investigates speech production in L2-immersed bilinguals, who may be affected by L1 attrition.

Research on L1 VOT in phonetic attrition is sparse, but there is broad evidence for L1 phonetic attrition at the segmental level (Bergmann et al., 2016; Chang, 2012; de Leeuw et al., 2013; Flege, 1987; Flege & Hillenbrand, 1984; Major, 1992; Mayr et al., 2012; Sancier & Fowler, 1997; Ulbrich & Ordin, 2014; Ventureyra, Pallier & Yoo, 2004) and at the suprasegmental level (de Leeuw et al., 2012; Mennen, 2004). L1 attrition affecting the segmental or suprasegmental level may surface as a global foreign accent (Bergmann et al., 2016; de Leeuw et al., 2010; Hopp & Schmid, 2013). Most of these studies on L1 phonetic attrition reported changes in the realization of L1 speech sounds or prosody under the influence of long term L2 use, and thus represent a context of language use that is similar to that of the participants in the present study (for short term L2 use, see Chang, 2012).

Phonetic attrition can surface as a drift of the L1 VOT values towards L2 VOT values. Four studies have observed phonetic attrition surfacing as durational changes in VOT in highly proficient L2 speakers (Flege, 1987; Major, 1992; Mayr et al., 2012; Sancier & Fowler, 1997). The bilinguals in these studies spoke Dutch, French or Portuguese, which have 'voiceless' short lag plosives, in addition to English, which has 'voiceless' aspirated plosives, like German. Native speakers of English produced shorter VOT in English 'voiceless' plosives when they frequently used French or Portuguese (Flege, 1987; Major, 1992). This was irrespective of whether they were immersed in the L2 or L1 context. Similarly, L1 speakers of French or Portuguese who were immersed in L2-English produced 'voiceless' plosives with longer VOT in L1-French and L1-Portuguese than the respective monolinguals (Flege, 1987; Sancier & Fowler, 1997). Further support for L1 phonetic attrition of VOT comes from a case study of a monozygotic twin who emigrated from the Netherlands to the United Kingdom 30 years before testing (Mayr et al., 2012). Her VOT production was evaluated against the speech of the other twin who lived in the Netherlands throughout her life. The emigrated twin exhibited longer – and therefore more English-like – VOT in 'voiceless' plosives than the Netherlands-based twin. By contrast, the emigrated twin's L1-Dutch 'voiced' plosives remained prevoiced and were thus not affected by L1 phonetic attrition. These four studies suggest that changes to the L1 VOT may be limited to bilinguals with high L2 proficiency, but appear to occur independently of the immersion context (Flege, 1987).

A more nuanced view on the role of the immersion context on durational changes to L1 VOT and target-like L2 VOT production is provided by longitudinal data of one Portuguese–English late bilingual (Sancier & Fowler, 1997). The speaker produced longer – and thus more English-like – VOT in L1-Portuguese and L2-English after several months of L2 immersion in the USA. In turn, the speaker produced shorter – and thus more Portuguese-like – VOT after subsequent L1 immersion in Brazil. These durational VOT changes were perceived by native listeners of Brazilian Portuguese who rated the speech as more accented right after the informant's stay in the USA than after a stay in Brazil. This study suggests that changes to L1 VOT do not necessarily reflect an irreversible loss of native-like L1 VOT.

Although L1 attrition surfacing as durational VOT changes has been observed in highly proficient L2 speakers (Flege, 1987; Major, 1992; Mayr et al., 2012; Sancier & Fowler, 1997), high L2 proficiency does not automatically lead to attrition of L1 VOT. Dutch L1 speakers who acquired native-like aspiration in L2-English maintained short lag VOT in Dutch ‘voiceless’ plosives (Simon, 2009; Simon & Leuschner, 2010). These speakers lived in their L1 country, which suggests that it may be more likely to maintain native-like L1 VOT with frequent native L1 input.

The observed cases of L1 VOT drift in ‘voiceless’ plosives are in line with the Speech Learning Model’s (SLM) assumed flexibility of L1 phonetic categories (Flege, 1995), and showed that L2 VOT can influence L1 VOT. This influence is not limited to an L2 immersion context, but rather seems related to frequency of language use. In addition, frequent L1 exposure through L1 immersion may help to prevent L1 attrition in highly proficient L2 speakers.

Only the case study of Mayr et al. (2012) included investigations of VOT in ‘voiced’ plosives, but found no evidence for phonetic attrition of L1 prevoicing. The present study follows up on this finding to address whether ‘voiced’ plosives are indeed resistant to durational changes of L1 VOT, while ‘voiceless’ plosives are frequently affected.

1.3 The current study

This study investigates VOT in the L1 and L2 speech of Dutch–German binational couples living in the Netherlands. Each couple consists of one partner with L1-Dutch and L2-German and one partner with L1-German and L2-Dutch. Within each couple, interactions in both languages are common as the two partners have at least one child that they raise bilingually. The L1-Dutch speakers are frequently exposed to German and to non-native Dutch at home through their German partner and their bilingual child or children. Similarly, the L1-German speakers are frequently exposed to Dutch and non-native German at home. The exposure to German in both groups of bilinguals is limited to the family context. Exposure to Dutch, by contrast, occurs in a variety of contexts and through multiple speakers.

In addition to a difference in immersion, the two groups face a different acquisition task: to produce target L2 VOT, the L1-Dutch speakers need to suppress

Dutch prevoicing and learn to produce German aspiration. The L1-German speakers need to suppress German aspiration and learn to produce Dutch prevoicing.

This study combines investigations of VOT in L2 acquisition and L1 attrition in both ‘voiceless’ and ‘voiced’ plosives in the same speakers. Addressing the speakers’ two languages and both voicing categories is essential to draw conclusions about the structure of bilinguals’ phonetic space and voicing systems. The use of bilingual couples as participants allows addressing L2 attainment by comparing one group of bilinguals’ L2 to the other group of bilinguals’ L1. This type of comparison offers two crucial advantages. First, a comparison between the L2 of one group of bilinguals and the L1 of the other group of bilinguals accounts for the characteristics of the speech to which the L2 speakers are daily exposed in their immediate social environment. Second, the L1 speech of bilinguals rather than monolinguals represents target speech that L2 speakers can in fact approach, as both groups’ phonologies encompass a similar number of phonemes.

The three questions we are specifically asking regarding both groups of bilinguals are whether both acquisition contexts allow to: (1) produce VOT differently in L1 and L2; (2) realize VOT in the L2 similarly to native speakers who are bilingual themselves; and (3) maintain L1 VOT that is similar to a monolingual control group consisting of speakers representative of the linguistic environment in which the participants acquired and used their L1 before they became bilingual.

Regarding the L1-Dutch speakers, we hypothesize that they produce longer than monolingual-like VOT in L1 ‘voiceless’ plosives, but maintain native-like prevoicing in L1 ‘voiced’ plosives (see Mayr et al., 2012). In L2-German, we expect the L1-Dutch speakers to produce ‘voiceless’ plosives with longer VOT than in Dutch, but with shorter VOT than the L1-German speakers. We further expect transfer of L1 prevoicing to L2 ‘voiced’ plosives.

Regarding the L1-German speakers, we expect to find shorter than monolingual-like VOT in L1 ‘voiceless’ plosives, and possibly prevoiced ‘voiced’ plosives to maintain a clear voicing contrast. If the L1-German speakers are indeed capable of producing prevoicing in L1-German and L2-Dutch, which has never been addressed in previous research, we expect them to be able to suppress aspiration and produce L2-Dutch ‘voiceless’ plosives with target-like short lag VOT.

2 Method

2.1 Participants

Ninety-seven speakers divided over four groups participated in this study: bilinguals with L1-Dutch and L2-German (*N*=18, 5 female), henceforth the L1D–L2G speakers; bilinguals with L1-German and L2-Dutch (*N*=23, 19 female), henceforth the L1G–L2D speakers; Dutch monolinguals (*N*=29; 26 female); and German monolinguals (*N*=27, 26 female). The participants were the parents of the preschoolers addressed in Stoehr et al. (2017) / Chapter 2. Table 1 provides detailed participant information.

Sixteen of the L1D–L2G speakers have had formal instruction to German in high school; the other two learned German only as adults when they met their German partner. The average age of first exposure to German of the L1D–L2G speakers was 13 years (range 1–28 years, *SD*=6). Regular exposure to German commenced for all L1D–L2G speakers when they met their German spouse in early adulthood. Further exposure to German now comes from their bilingual child or children. Twelve L1D–L2G speakers reported frequent use of Dutch and German. Six reported frequent use of Dutch and occasional use of German.

The L1G–L2D speakers learned Dutch at an average age of 23 years (range 8–33 years, *SD*=6), when they moved to the Netherlands. One participant learned Dutch at school before she was regularly exposed to Dutch through her partner. Twenty-two of the participants in this group reported frequent use of German and Dutch. One participant reported frequent use of German and occasional use of Dutch.

Although not all participants reported knowledge of an additional language besides Dutch and German, schooling in the Netherlands and Germany requires all students to study English. Language teachers in these countries traditionally are non-native speakers of English, so that English instruction does not equal exposure to native English accents.

The majority of the bilingual participants consisted of 17 Dutch–German binational couples, contributing one partner to the L1D–L2G group and the other partner to the L1G–L2D group. One additional participant in the L1D–L2G group and six participants in the L1G–L2D group participated without their partners. The bilinguals were tested in different provinces across the Netherlands.

Of the Dutch monolinguals, two reported some knowledge of German, and three reported speaking English sporadically. All Dutch monolinguals were tested in or around Nijmegen in the Central Eastern Netherlands. Four of the monolingual

Table 1. Participant overview.

ID	L1	Gender	Frequent German	Frequent Dutch	AoA L2	Dutch at work	L2 active	L2 passive	Other L2*
L1-G-01	German	F	✓	✓	20	✓	4	4	
L1-G-02	German	M	✓	✓	13	✗	5	5	
L1-G-03	German	M	✓	✓	?	✓	4	4	
L1-G-04	German	F	✓	?	31	✓	4	5	
L1-G-06	German	F	✓	✓	23	✗	3	4	
L1-G-07	German	F	✓	✓	20	✓	4	4	
L1-G-10	German	F	✓	✓	24	✓	5	5	
L1-G-12	German	F	✓	✓	20	✓	3	4	
L1-G-13	German	F	✓	✓	25	✓	4	4	
L1-G-15	German	F	✓	✓	20	✓	5	5	
L1-G-16	German	F	✓	✓	8	✓	5	5	F
L1-G-17	German	F	✓	✓	25	✗	3	4	
L1-G-18	German	F	✓	✗	27	✗	4	5	
L1-G-19	German	F	✓	✓	23	✓	5	5	
L1-G-21	German	F	✓	✓	25	✓	4	4	
L1-G-23	German	F	✓	✓	33	✓	4	5	
L1-G-24	German	F	✓	✓	30	✓	4	5	F
L1-G-26	German	F	✓	✓	25	✓	4	4	DA, P, N
L1-G-27	German	F	✓	✓	20	✓	4	4	
L1-G-29	German	M	✓	✓	16	✓	5	5	
L1-G-31	German	M	✓	✓	23	✓	4	4	
L1-G-32	German	F	✓	✓	19	✓	3	4	
L1-G-33	German	F	✓	✓	33	✗	4	4	
L1-D-02	Dutch	F	✓	✓	13	✓	4	4	
L1-D-03	Dutch	F	✓	✓	?	✓	4	4	
L1-D-06	Dutch	M	✓	✓	12	✓	3	4	
L1-D-07	Dutch	M	✓	✓	14	✓	3	3	
L1-D-10	Dutch	M	✗	✓	14	✓	4	4	
L1-D-11	Dutch	F	✓	✓	28	✓	4	4	
L1-D-12	Dutch	M	✗	✓	14	✓	2	2	
L1-D-16	Dutch	M	✓	✓	12	✓	3	4	I, DA
L1-D-18	Dutch	M	✗	✓	13	✓	3	4	
L1-D-19	Dutch	M	✓	✓	13	✓	3	3	
L1-D-21	Dutch	M	✓	✓	1	✓	4	4	
L1-D-24	Dutch	M	✓	✓	12	✓	4	4	
L1-D-26	Dutch	M	✗	✓	12	✓	4	4	
L1-D-27	Dutch	M	✓	✓	6	✓	3	3	
L1-D-29	Dutch	F	✓	✓	13	✓	3	3	
L1-D-31	Dutch	F	✗	✓	25	✓	3	3	
L1-D-32	Dutch	M	✓	✓	13	✓	4	4	
L1-D-33	Dutch	M	✗	✓	14	✓	2	3	

*All speakers had English instruction during high school. Codes: ✓=yes, ✗=no, ?=no information provided. Additional L2: DA=Danish, F=French, I=Italian, P=Portuguese, N=Norwegian. L2 active: 5=native fluency, 4=very fluent, 3=quite fluent, 2=somewhat fluent, 1=limited fluency, 0=virtually no fluency. L2 passive: 5=native understanding, 4=excellent understanding, 3=good understanding, 2= some understanding, 1= limited understanding, 0=almost no understanding.

German participants had some knowledge of Dutch, but none of them reported regular use of a language different from German. The German monolinguals were tested in Central Western Germany (*N*=27) and Northern Germany (*N*=2). Like the bilinguals, all monolinguals had studied English in high school.

2.2 Materials and procedure

The elicited target words were the same as described in Chapter 2, section 2.2.1. Testing took place in a quiet room in the participants' homes, after the participants signed informed consent for their family to participate in the study. When both participants of a couple completed the task during the same testing session, the other participant left the room during the recordings. The participants were shown pictures of the target words and they were asked to name them at a comfortable pace without using a determiner. The participants then filled out a language background questionnaire, while their children completed three tasks for the studies reported in chapters 2, 3, and 4. Finally, the participants named the pictures in their other language. The language order was counterbalanced across participants. The picture naming task took approximately three minutes per language.

2.3 Recordings and VOT measurements

See Chapter 2, section 2.3 for a description of the recordings and measurements. Inter-coder reliability based on 25% of the data reported in this chapter indicated 99% agreement.

3 Results

In this section, we first provide an overview of the descriptive statistics of 'voiceless' plosives (Table 2 and Figure 1) and 'voiced' plosives (Tables 3 and 4, Figure 2). We then present the statistical models (Table 5) before we turn to the statistical effects of Language and Language Background on VOT, which are summarized in Table 6.

Table 2 provides the means and standard deviations of VOT per 'voiceless' plosive over participants by language and language background. Both groups of bilinguals produced overall longer VOT in German than in Dutch. In each language, the bilinguals produced L1 VOT intermediate to the monolinguals' L1 VOT and the L2 VOT of the other group of bilinguals. In Dutch, the L1D–L2G speakers produced minimally longer VOT than the monolinguals, and shorter VOT than the L1G–L2D speakers. In German, the L1G–L2D speakers produced VOT that was intermediate to the monolinguals' overall longer VOT and the L1D–L2G speakers overall shorter VOT. Figure 1 visualizes these findings by consonantal place of articulation.

VOT of 'voiced' plosives was bimodally distributed in 47 of the 70 participants in Dutch and in 51 of the 68 participants in German. VOT of 'voiced' plosives was therefore treated categorically as either prevoiced (negative VOT) or short lag (short positive VOT). Table 3 shows the mean percentages and standard deviations of the 'voiced' plosives produced with prevoicing (and inversely related short lag VOT) over participants together with the total number of analyzable prevoiced and short lag tokens per 'voiced' plosive by language and language background. Both groups of bilinguals produced overall more prevoiced tokens in Dutch than in German, although this difference is more pronounced in the L1G–L2D speakers. In Dutch, the L1D–L2G speakers produced the highest percentage of 'voiced' plosives with prevoicing, closely followed by the monolingual Dutch speakers. This small between-group difference may be ascribed to the larger number of males in the L1D–L2G group, who typically produce more prevoicing than females (Ryalls et al., 1997). The L1G–L2D speakers produced a lower percentage of prevoiced plosives in Dutch than the two groups of Dutch native speakers. In German, the monolinguals produced the lowest percentage of prevoiced plosives, followed by the L1G–L2D speakers. The L1D–L2G speakers produced the highest percentage of prevoiced plosives. Figure 2 visualizes the percentages of prevoiced plosives by language and consonantal place of articulation across the groups. The devoiced 'voiced' plosives had VOT values close to 10 ms in both languages and all groups (Table 4).

Table 2. VOT in ms by place of articulation and group over participants.

		Dutch			German		
		L1G–L2D	L1D–L2G	MonoDutch	L1D–L2G	L1G–L2D	MonoGerman
	M	21	10	8	23	38	45
/p/	*SD*	15	6	5	19	17	18
	Total #	147	105	173	109	140	159
	M	31	23	21	48	59	69
/t/	*SD*	13	9	10	20	19	17
	Total #	141	111	179	108	140	169
	M	43	31	28	44	58	72
/k/	*SD*	16	13	10	18	18	20
	Total #	139	110	171	112	140	165
	Overall *M*	32	21	19	38	52	62

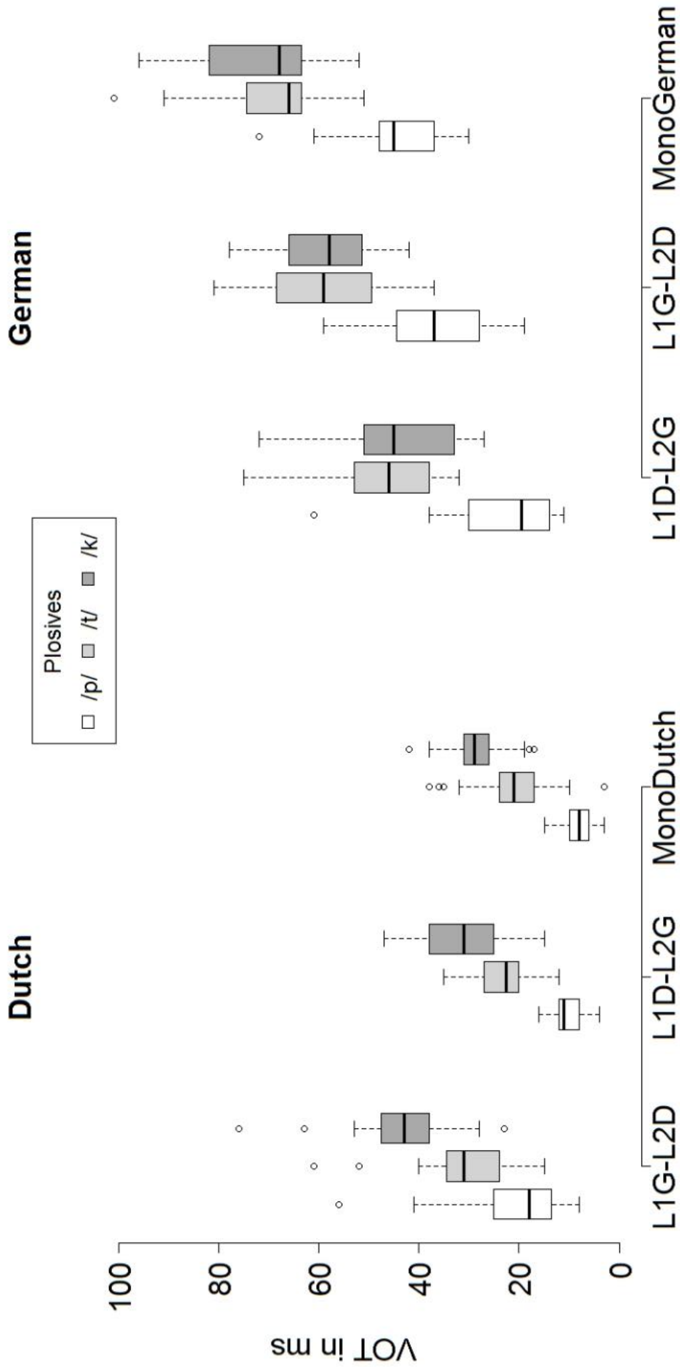


Figure 1. Voice onset time (VOT) of ‘voiceless’ plosives by language background over participants.

Notes. L1G–L2D=bilinguals with German as first language and Dutch as second language; L1D–L2G=bilinguals with Dutch as first language and German as second language; MonoDutch=Dutch monolinguals; MonoGerman=German monolinguals.

5

Table 3. Mean percentage of prevoiced plosives by place of articulation and group over participants.

		Dutch			German		
		L1G–L2D	L1D–L2G	MonoDutch	L1D–L2G	L1G–L2D	MonoGerman
	M % prevoiced	66	91	87	87	38	26
/b/	*SD*	35	24	22	20	37	34
	Total #	95/143	96/106	149/172	93/107	53/140	42/158
	M % prevoiced	64	82	79	64	26	22
/d/	*SD*	33	22	24	30	31	24
	Total #	86/139	93/110	133/165	66/103	37/145	38/177
	Overall *M* % prevoiced	65	87	83	76	32	24

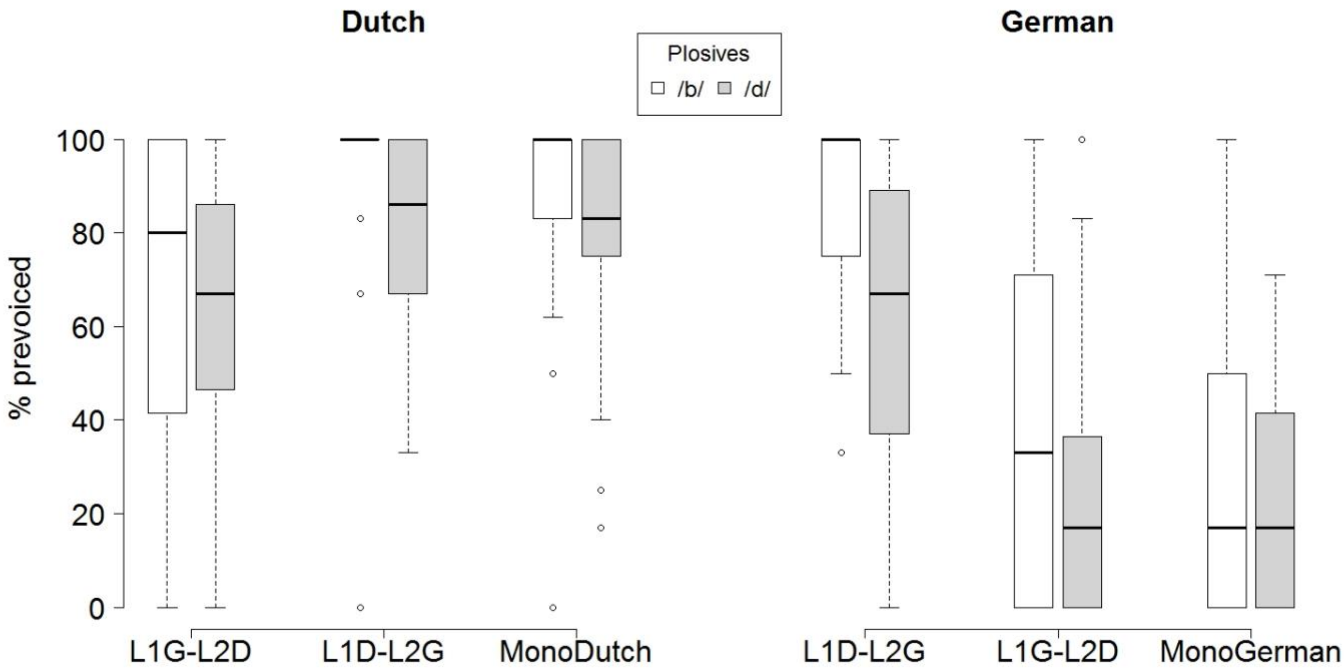


Figure 2. Percentage of ‘voiced’ plosives produced with prevoicing by language background over participants.

Notes. L1G–L2D=bilinguals with German as first language and Dutch as second language; L1D–L2G=bilinguals with Dutch as first language and German as second language; MonoDutch=Dutch monolinguals; MonoGerman=German monolinguals.

5

Table 4. VOT in ms of short lag 'voiced' plosives by place of articulation and group over participants.

		Dutch			German		
		L1G–L2D	L1D–L2G	MonoDutch	L1D–L2G	L1G–L2D	MonoGerman
/b/	*M*	9	11	5	8	7	6
	SD	3	2	2	2	3	3
	Total #	48	10	23	14	87	116
/d/	*M*	12	14	13	13	12	12
	SD	7	3	9	5	4	4
	Total #	53	17	32	37	108	139
	Overall *M*	11	13	9	11	10	9

3.1 Description of the statistical models

Statistical analyses using mixed effects regression were performed in R (R Core Team, 2013). An α-level of .05 was adopted throughout. VOT of the 'voiceless' plosives /p/, /t/ and /k/ was analyzed as a continuous variable using mixed effects linear regression. VOT of the 'voiced' plosives /b/ and /d/ was analyzed as a categorical variable using mixed effects logistic regression to address the aforementioned bimodal distribution of VOT. Negative VOT values were coded as 'prevoiced' and values equal to or greater than zero were coded as 'short lag'. Due to the use of different regression types, each research question was addressed with separate models for 'voiceless' and 'voiced' plosives. Each research question was furthermore addressed with specific between-group or within-group comparisons, which are outlined below.

The bilinguals' differentiation of L1 and L2 VOT was assessed with within-group comparisons of the bilinguals' Dutch and German. This L1–L2 comparison was conducted separately for the L1G–L2D speakers and the L1D–L2G speakers, and the independent variable (IV) of main interest was *Language* (Dutch vs. German).

Two between-group analyses addressed nativelikeness of the bilinguals' VOT in the two languages. L2 attainment was assessed by comparing the bilinguals' L2 VOT to the other bilinguals' L1 VOT. L1 attrition was assessed by comparing the bilinguals' L1 VOT to the VOT of an independent sample of monolinguals. The IV of main interest in all between-group analyses was *Language Background* (the bilinguals' L2 vs. the other bilinguals' L1; the bilinguals' L1 vs. the monolinguals' L1).

Additional IVs were used in all models to account for item-related and participant-related variance due to factors that are known to impact on VOT. Item-related IVs for analyses on 'voiceless' plosives were *Place of Articulation* of the plosive (/p/ vs. /t/ and /t/ vs. /k/) and *Word Length* (monosyllabic vs. disyllabic). The item-related IV for analyses on 'voiced' plosives was *Place of Articulation* (/b/ vs. /d/). The participant-related IV in all analyses was *Gender*.

Table 5 provides an overview of the model specifications for each group comparison. All models comprised interactions between the IV of main interest and the other IVs, except for the models on L2 attainment, where simplification due to model convergence problems was required. Significant interactions were explored in separate follow-up analyses for each level of the IVs.

Table 5. Model specifications.

Groups	Analysis	Fixed effects	Interactions	Random effects & intercept	Nesting	Random slopes
Bilingual L1 vs. bilingual L2 (L1G–L2D speakers & L1D–L2G speakers)	voiceless (linear)	Language		Participant		Language
		Gender	Language*Gender			PoA-LC
		PoA-LC	Language*PoA-LC			PoA-CD
		PoA-CD	Language*PoA-CD			Word Length
		Word Length	Language*Word Length	Item		none
	voiced (logistic)	Language		Participant		Language
		Gender	Language*Gender			PoA
		PoA	Language*PoA	Item		none
Bilingual L2 vs. bilingual native speakers (Dutch & German)	voiceless (linear)	LangBackgr.		Participant	Couple	PoA-LC
		Gender	LangBackgr.*Gender			PoA-CD
		PoA-LC	LangBackgr.*PoA-LC			Word Length
		PoA-CD	LangBackgr.*PoA-CD[i]	Item		LangBackgr.
		Word Length	LangBackgr.*WordLLength[ii]			
	voiced (logistic)	LangBackgr.		Participant	Couple	PoA
		Gender	LangBackgr.*Gender	Item		LangBackgr.
		PoA	LangBackgr.*PoA			

Table 5. Model specifications (continued).

Groups	Analysis	Fixed effects	Interactions	Random effects & intercept	Nesting	Random slopes
Bilingual L1 vs. monolingual native speakers (Dutch & German)	voiceless (linear)	LangBackgr.		Participant		PoA-LC
		Gender	LangBackgr.*Gender			PoA-CD
		PoA-LC	LangBackgr.*PoA-LC			Word Length
		PoA-CD	LangBackgr.*PoA-CD	Item		LangBackgr.
		Word Length	LangBackgr.*Word Length			
	voiced (logistic)	LangBackgr.		Participant		PoA
		Gender	LangBackgr.*Gender	Item		LangBackgr.
		PoA	LangBackgr.*PoA			

Notes: LangBackgr.=Language Background; PoA-LC=Place of Articulation: Labial vs. Coronal; PoA-CD=Place of Articulation: Coronal vs. Dorsal; [i] only in Dutch model due to convergence problems; [ii] only in German model due to convergence problems.

5

3.2 Results of the statistical models

This section presents the main findings of the three research questions. The first two analyses addressed the bilinguals' differentiation of VOT in the L1 and L2. Subsequent analyses addressed the bilinguals' L2 attainment and potential L1 attrition. Lastly, we present findings on variability specific to the target words and participants that did not contribute to the main results.

a) Differentiation between L1 and L2 VOT within the bilinguals. The analyses on language differentiation in the L1G–L2D speakers showed that they produced VOT differently when speaking German compared to when speaking Dutch. The L1G–L2D speakers specifically produced longer VOT in 'voiceless' plosives when speaking German ($\beta=16.22$, $SE=2.41$, $t=6.72$, $p<.001$), and a higher percentage of 'voiced' plosives with prevoicing when speaking Dutch ($\beta=0.95$, $SE=0.34$, $z=2.84$, $p<.005$). In addition, an interaction between *Language* and *Place of Articulation* ($\beta=-6.37$, $SE=2.87$, $t=-2.22$, $p=.026$) revealed that the L1G–L2D speakers produced longer VOT in /k/ than in /t/ in Dutch ($\beta=12.31$, $SE=3.32$, $t=3.70$, $p<.001$), but not in German ($\beta=-0.45$, $SE=4.91$, $t=-0.09$, $p>.250$).

The L1D–L2G speakers produced distinct VOT for Dutch and German 'voiceless' plosives, but not for 'voiced' plosives. They produced 'voiceless' plosives with longer VOT in German than in Dutch ($\beta=13.83$, $SE=2.44$, $t=5.68$, $p<.001$), but no difference in the percentage of 'voiced' plosives produced with prevoicing in Dutch and in German was detected ($\beta=0.43$, $SE=0.28$, $z=1.54$, $p=.124$). An interaction between *Language* and *Word Length* ($\beta=2.60$, $SE=1.25$, $t=2.07$, $p=.038$) revealed that the L1D–L2G speakers produced 'voiceless' plosives with longer VOT in monosyllabic than in disyllabic words in German ($\beta=4.62$, $SE=2.05$, $t=2.25$, $p=.024$), but not in Dutch ($\beta=-0.39$, $SE=0.98$, $t=-0.40$, $p>.250$). Overall, the results on phonetic differentiation between L1 and L2 suggest that Dutch–German late bilinguals produced VOT differently in L1 and L2 with the exception of the L1D–L2G speakers' production of 'voiced' plosives.

b) L2 attainment and L1 attrition. The following four analyses concerned the bilinguals' VOT production in both their L2 and their L1. The reference point for L2 attainment was the other bilinguals' L1. The reference point for L1 attrition was the speech of monolingual native-speakers.

L1G–L2D speakers. The analyses on L2 attainment in the L1G–L2D speakers showed that they attained native-like VOT in L2-Dutch for /p/ and /t/, but not for /k/ or 'voiced' plosives. In L2-Dutch 'voiceless' plosives, no overall VOT differences were detected between the L1G–L2D speakers and the L1D–L2G speakers ($\beta=-2.10$, $SE=1.45$, $t=-1.45$, $p=.147$), but an interaction between *Language Background* and *Place of Articulation* ($\beta=-2.30$, $SE=1.13$, $t=-2.04$, $p=.041$) revealed that the L1G–L2D speakers in fact produced longer VOT in /k/ than the L1D–L2G speakers ($\beta=-4.91$, $SE=1.68$, $t=-2.92$, $p=.004$). In L2-Dutch 'voiced' plosives, the L1G–L2D produced a lower percentage of prevoiced plosives than native speakers ($\beta=-0.95$, $SE=0.46$, $z=-2.06$, $p=.039$)[16].

The analyses on L1 attrition in the L1G–L2D speakers showed that their L1-German VOT of 'voiceless' but not 'voiced' plosives is affected by L1 attrition. The L1G–L2D speakers produced L1-German 'voiceless' plosives with shorter VOT than monolinguals ($\beta=-6.94$, $SE=3.10$, $t=-2.24$, $p=.025$). By contrast, no differences in the percentage of prevoicing between the L1G–L2D speakers and monolinguals were observed ($\beta=-0.13$, $SE=0.50$, $z=-0.25$, $p>.250$).

L1D–L2G speakers. The analyses on L2 attainment in the L1D–L2G speakers showed that they produced non-native VOT in L2-German. The L1D–L2G speakers produced L2-German 'voiceless' plosives with shorter VOT than the L1G–L2D speakers ($\beta=-6.57$, $SE=1.65$, $t=-3.97$, $p<.001$). Similarly, they produced a higher percentage of German 'voiced' plosives with prevoicing than the L1G–L2D speakers ($\beta=-1.06$, $SE=0.28$, $z=-3.79$, $p<.001$). An interaction between *Language Background* and *Gender* ($\beta=-0.92$, $SE=0.37$, $z=-2.49$, $p=.013$) did not reveal any gender-related differences in the production of prevoicing in the L1D–L2G group ($\beta=-0.50$, $SE=0.41$, $z=-1.20$, $p=.230$), but rather revealed that males in the L1G–L2D group produced a higher percentage of prevoiced 'voiced' plosives than females ($\beta=1.67$, $SE=0.51$, $z=3.30$, $p<.001$).

The analyses on L1 attrition in the L1D–L2G speakers did not find evidence for attrition of L1-Dutch VOT. The L1D–L2G speakers neither produced L1-Dutch 'voiceless' plosives ($\beta=1.86$, $SE=1.16$, $t=1.60$, $p=.110$) nor 'voiced' plosives

[16] A potential caveat in the comparison of prevoicing in Dutch between the L1G–L2D speakers and the L1D–L2G speakers is that the latter group comprises a higher number of males. The higher percentage of prevoicing in the L1D–L2G speakers could thus be ascribed to the gender difference rather than to the language background (Ryalls et al., 1997), even though *Gender* is taken into account in the model. To further investigate whether the L1G–L2D speakers prevoiced less in Dutch than native speakers, a second model was run in which the Dutch monolinguals served as reference. This model showed that the differences in the percentage of prevoicing between the L1G–L2D speakers and Dutch monolinguals was significant in a one-tailed, but not a two-tailed comparison ($\beta=0.77$, $SE=0.43$, $z=1.80$, $p=.073$). The effect of *Language Background* on 'voiced' plosives is interpreted in the following section, but the reader is asked to recall that the effect is small to marginal, depending on the reference group.

(β=−0.06, SE=0.44, z=−0.13, p>.250) detectably different from Dutch monolinguals. In sum, the results of the analyses on L2 attainment and L1 attrition show that only the L1G–L2D bilinguals who were immersed in the L2 country partially attained native-like L2 VOT. Similarly, only the L1D–L2G bilinguals who were immersed in the L1 country maintained native-like L1 VOT.

c) *Variability related to the words and participants*. In the following, we present the significant findings on the IVs relating to the target words and participants. As the bilinguals were part of three analyses, the results of an IV for a group was considered significant when at least one analysis including the group yielded significance for an IV. The complete model output of all models is provided in the Appendix.

In analyses on 'voiceless' plosives, all groups produced shorter VOT for /p/ than for /t/ in Dutch and in German, and all groups produced longer VOT for /k/ than for /t/ only in Dutch, but not in German. In addition, all groups produced longer VOT in monosyllabic than in disyllabic words in German, but not in Dutch. In analyses on 'voiced' plosives, all groups prevoiced /b/ more frequently than /d/ in both languages. In all groups except the Dutch monolinguals, males prevoiced more frequently than females. Late bilinguals thus produce language-specific within-category VOT variability related to consonantal place of articulation and word length.

4 Summary

The present study investigated how two groups of Dutch–German late bilinguals in the Netherlands realize the voicing contrast in both Dutch and German by means of voice onset time (VOT). The bilinguals who speak Dutch as native language and German as the L2 are referred to as L1D–L2G speakers, and the bilinguals who speak German as native language and Dutch as the L2 are referred to as L1G–L2D speakers. To achieve native-like L2 VOT, the L1D–L2G speakers need to acquire aspiration for L2-German 'voiceless' plosives and suppress prevoicing for L2-German 'voiced' plosives. The L1G–L2D speakers need to suppress aspiration in L2-Dutch 'voiceless' plosives and consistently prevoice L2-Dutch 'voiced' plosives. We investigated whether (1) both groups of bilinguals produced VOT differently in L1 and L2; (2) both groups of bilinguals achieved native-like L2 VOT; and (3) both groups of bilinguals maintained native-like L1 VOT.

Table 6. Results overview.

				Language	**Language Background**
RQ 1	Bilingual Dutch vs. bilingual German	L1G–L2D speakers	voiceless	Longer VOT in German ***	--
			voiced	Higher percentage of prevoicing in Dutch **	--
		L1D–L2G speakers	voiceless	Longer VOT in German ***	--
			voiced	n.s.	--
RQ 2	Bilingual L2-Dutch vs. bilingual L1-Dutch	L1G–L2D speakers	voiceless	--	n.s.
			voiced	--	L2 speakers: lower percentage of prevoicing than L1 speakers *
	Bilingual L2-German vs. bilingual L1-German	L1D–L2G speakers	voiceless	--	L2 speakers: shorter VOT than L1 speakers ***
			voiced	--	L2 speakers: higher percentage of prevoicing than L1 speakers ***
RQ 3	Bilingual L1-German vs. monolingual German	L1G–L2D speakers	voiceless	--	Bilingual L1 speakers: shorter VOT than monolinguals *
			voiced	--	n.s.
	Bilingual L1-Dutch vs. monolingual Dutch	L1D–L2G speakers	voiceless	--	n.s.
			voiced	--	n.s.

*** p<.001; ** p<.01; * p<.05; n.s. p>.05

5

The L1G–L2D speakers produced ‘voiceless’ plosives with short lag VOT in L2-Dutch /p/ (*M*=21 ms) and /t/ (*M*=31 ms), and slight aspiration in Dutch /k/ (*M*=43 ms), while they aspirated L1-German ‘voiceless’ plosives (*M*=52 ms). Similarly, the L1G–L2D speakers prevoiced a higher percentage of ‘voiced’ plosives in L2-Dutch (65%) than in L1-German (32%). The L1G–L2D speakers produced the remaining ‘voiced’ plosives with short lag VOT that was virtually alike in L2-Dutch (*M*=11 ms) and L1-German (*M*=10 ms), and considerably shorter than their VOT of L2-Dutch ‘voiceless’ plosives (*M*=32 ms). However, the L1G–L2D speakers did not acquire new VOT ranges, as aspiration, short lag and prevoicing are all observed in the speech of monolingual German-speaking adults as well. Instead, the acquisition task they accomplished was redefining their phonetic space: in addition to the pre-existing aspirated category (German /p/, /t/, /k/), the L1G–L2D speakers restructured their ‘prevoicing to short lag’ phonetic space into three individual categories: short lag > 20 ms (Dutch /p/, /t/, /k/), short lag ~10 ms (German /b/, /d/ and sometimes Dutch /b/, /d/), and prevoicing (Dutch /b/, /d/ and sometimes German /b/, /d/). This L1-German–L2-Dutch phonetic system displays absolute phonological differentiation between ‘voiceless’ and ‘voiced’ plosives, as well as absolute by-language differentiation between Dutch and German ‘voiceless’ plosives, but gradient by-language differentiation between Dutch and German ‘voiced’ plosives.

The L1G–L2D speakers seem to have attained native-like Dutch short lag VOT, at least for /p/ and /t/, but they did not yet reach native-like consistent prevoicing. In German, their VOT partly seems to be affected by language attrition, as revealed by shorter than monolingual-like VOT in ‘voiceless’ plosives. ‘voiced’ plosives, by contrast, seem to remain unaffected by language attrition.

The L1D–L2G speakers produced ‘voiceless’ plosives with longer VOT in L2-German (*M*=38 ms) than in L1-Dutch (*M*=21 ms), but they prevoiced the majority of ‘voiced’ plosives in both L2-German (76%) and L1-Dutch (87%). The L1D–L2G speakers seem to have three phonetic categories: a new L2 long lag category ~40 ms (German /p/, /t/, /k/), their pre-existing L1 short lag category ~20 ms (Dutch /p/, /t/, /k/), and a prevoiced category that merges L2 with L1 ‘voiced’ plosives (Dutch and German /b/, /d/). Their L1-Dutch–L2-German phonetic space displays absolute phonological differentiation between ‘voiceless’ and ‘voiced’ plosives, whereas by-language differentiation between Dutch and German is present for ‘voiceless’ plosives, but absent for ‘voiced’ plosives.

The L1D–L2G speakers’ differentiation between ‘voiceless’ plosives between Dutch and German does not go hand in hand with attainment of native-like VOT in German. They hardly aspirate /p/ (*M*=23 ms) and produce less aspiration in

/t/ (*M*=48 ms) and /k/ (*M*=44 ms) than the L1G–L2D speakers. Similarly, they prevoiced a higher percentage of ‘voiced’ plosives in L2-German (76%) compared to the L1G–L2D speakers (32%). Despite the L1D–L2G speakers’ exposure to German at home, their Dutch VOT was not affected by attrition and remained similar to that of monolingual native speakers of Dutch.

5 Discussion

In the following, we first interpret the results in light of the Speech Learning Model’s (SLM) equivalence classification and contrast maintenance hypotheses (Flege, 1995). We then discuss immersion and language use, articulatory constraints, and foreign accentedness as additional explanations of the results.

5.1 Equivalence classification and contrast maintenance

The SLM (Flege, 1995) attempts to explain L2 phonetic attainment in relation to the L1 phonetic system. The two main concepts applicable to this study are equivalence classification and contrast maintenance. Differential acquisition, that is deviation from native norms, was observed in the L1D–L2G speakers for both L2-German ‘voiceless’ and ‘voiced’ plosives, and in the L1G–L2D speakers for L2-Dutch ‘voiceless’ /k/ and ‘voiced’ plosives.

One account within the SLM to explain such differential acquisition is equivalence classification (Flege, 1987, 1995): L2 speakers perceive L2 sounds into their preexisting L1 categories, and thus produce them in line with their L1 categories. However, equivalence classification cannot explain the specific patterns of differential acquisition in the present results. The L1G–L2D speakers prevoiced less frequently in Dutch than native speakers, but they prevoiced more frequently in L2-Dutch than in L1-German. Similarly, the L1D–L2G speakers did not produce native-like aspiration in L2-German, but they produced ‘voiceless’ plosives with longer VOT in L2-German than in L1-Dutch. The observed differences between Dutch and German in the L1G–L2D speakers and the L1D–L2G speakers indicate that they perceive differences between the respective Dutch and German plosives. An alternative account for the differential acquisition of Dutch prevoicing and German aspiration lies in articulatory constraints, as discussed in detail below.

Equivalence classification has further limitations explaining the L1D–L2G speakers’ transfer of prevoicing from L1-Dutch to L2-German. Prevoicing is the main cue for Dutch native listeners’ voicing perception (van Alphen & Smits, 2004).

Equivalence classification would thus predict that the L1D–L2G speakers perceive German short lag plosives into their equivalent Dutch short lag ‘voiceless’ category and thus produce German ‘voiced’ plosives without any prevoicing. The need to maintain contrast between L2-German ‘voiceless’ and ‘voiced’ plosives offers an alternative explanation for the L1D–L2G speakers transfer of prevoicing to German.

Contrast maintenance is a second hypothesis within the SLM to explain differential L2 phonetic acquisition, and suggests acquisition of deviating phonetic categories in L2 to maintain contrast with already existing phonetic categories. The L1D–L2G speakers may need to produce prevoicing in L2-German to maintain a distinction between their ‘voiced’ and ‘voiceless’ categories. The VOT of their German ‘voiceless’ plosives, especially in /p/, is perhaps too short to be contrasted with target-like short lag ‘voiced’ plosives (Flege & Eefting, 1987a; Keating, 1984).

In contrast to the SLM’s predictions of differential acquisition, the L1G–L2D speakers reached native-like VOT in L2-Dutch /p/ and /t/. Their short lag space was initially occupied by L1-German ‘voiced’ plosives, and therefore acquiring L2-Dutch short lag ‘voiceless’ plosives constitutes an intricate task: keeping L2-Dutch ‘voiceless’ short lag plosives separate from L1-German ‘voiced’ short lag plosives requires restructuring of L1 phonetic categories. Native-like L2 phonetic categories can thus be acquired under favorable conditions, including long-term L2 immersion with diverse L2 use, simple articulatory gestures, and the social need to reduce a potential foreign accent. The effect of these conditions on L2 attainment and L1 attrition is discussed in detail below.

5.2 Immersion and language use

The two investigated immersion contexts, full immersion in an L2 environment and immersion in the L2 at home, are comparable in that both contexts involve natural and frequent use of the L2. Full L2 immersion is inherently tied with L2 use in a variety of contexts and also with numerous speakers, whereas it largely limits L1 use to conversations within the family. By contrast, L2 immersion at home limits L2 use to interactions within the family, while the L1 is continuously used outside the home in a variety of contexts and with numerous speakers. Successful L2 acquisition as well as L1 attrition seem to be limited to an immersion context that involves drastic reduction of native L1 contact due to extensive L2 use, as is the case for the L1G–L2D speakers.

One aspect of full immersion that may influence the outcomes of L2 acquisition is exposure to multiple speakers, which is beneficial in monolingual and heritage L1 acquisition (Gollan et al., 2015; Seidl et al., 2014). Such diverse L2 exposure was experienced by the L1G–L2D speakers (exposed to Dutch in and outside the home), who acquired target L2-Dutch 'voiceless' plosives, but not by the L1D–L2G speakers (exposed to German in the home only) who did not acquire target L2-German plosives.

Conversely, frequent L1 contact and use in diverse contexts and with multiple speakers may be necessary to prevent phonetic L1 attrition, as has previously been suggested by Mayr et al. (2012). This hypothesis is in line with previous research that found quality and quantity of native language input to play a crucial role in L1 maintenance (de Leeuw et al., 2010). Only the L1D–L2G speakers, who were exposed to L1-Dutch outside the home, maintain native-like L1 VOT. Without frequent and diverse exposure to the L1, the more prominent L2 is likely to impact on the L1 phonetic categories. The L1G–L2D speakers, whose L1-German use was limited to the family context, were affected by L1 phonetic attrition surfacing as shorter than native-like aspiration in L1-German 'voiceless' plosives. Diversity of language use and exposure are important topics for future research into the circumstances that lead to successful L2 acquisition and L1 maintenance.

5.3 Articulatory constraints

Articulatory constraints seem to be at play when it comes to successful L2 acquisition and L1 maintenance of VOT. In comparison to short lag VOT, aspiration requires an additional timing component, as the glottis must remain open during burst release and be closed shortly after. Prevoicing requires complete glottal closure, and initiation and sustainment of vocal fold vibration before burst release (Kewley-Port & Preston, 1974).

Articulatory least complex short lag VOT was successfully acquired for L2-Dutch /p/ and /t/ by the L1G–L2D bilinguals. L1 short lag VOT was furthermore successfully maintained by the L1D–L2G speakers for L1-Dutch 'voiceless' plosives and also by the L1G–L2D speakers for L1-German 'voiced' plosives. Despite the articulatory simplicity of short lag VOT, it is still remarkable that the L1G–L2D speakers were able to suppress their L1-German aspiration and produce short lag VOT in /p/ and /t/ in L2-Dutch. To our knowledge, such suppression of aspiration in an L2 with target short lag 'voiceless' plosives has never been reported in late L2 learners, and instead aspiration was carried over from L1 to L2 (Flege, 1987).

Although short lag VOT is allegedly easy to produce (Kewley-Port & Preston, 1974), the L1D–L2G speakers produced L2-German ‘voiced’ plosives with prevoicing instead of short lag VOT. As discussed in section 5.1, the production of prevoiced ‘voiced’ plosives in L2-German may be caused by the need to maintain phonetic contrast with the L2-German ‘voiceless’ plosives, which were produced with shorter than target-like VOT.

Articulatory more complex aspiration was not completely acquired by the L1D–L2G speakers in L2-German. Similarly, the target aspirated L1-German ‘voiceless’ plosives of the L1G–L2D speakers appear to be affected by phonetic attrition.

The articulatory most complex Dutch prevoicing was not completely acquired by the L1G–L2D speakers, but was successfully maintained by the L1D–L2G speakers. Despite the complex velopharyngeal activity involved in the production of prevoicing, the L1G–L2D speakers, and also the German monolinguals, are well capable of initiating velopharyngeal adjustments to close the glottis prior to oral release of the consonant, as evidenced by occasional occurrences of prevoicing in their speech. They may, however, not necessarily be able to control the required muscular activities to a similar extent as native speakers of a prevoicing language, which results in overall fewer productions of prevoicing in their speech.

5.4 Foreign accent

Another factor contributing to successful L2 acquisition and L1 maintenance may be accentedness and the associated social stigmatization (Fuertes, Gottdiener, Martin, Gilbert & Giles, 2012; Kinzler, Dupoux & Spelke, 2007). Production of aspiration in a language without aspiration, such as Dutch, is associated with a foreign accent (Flege, 1984; Major, 1987; Riney & Takagi, 1999; Sancier & Fowler, 1997; Schoonmaker-Gates, 2015). Dutch short lag ‘voiceless’ plosives were successfully acquired by the L1G–L2D speakers and maintained by the L1D–L2G speakers. The social need to avoid stigmatization may be advantageous for the suppression of aspiration in L2-Dutch and the maintenance of short lag VOT in L1-Dutch.

Not all non-native VOT productions are associated with a perceived foreign accent: when target short lag ‘voiced’ plosives are prevoiced, listeners do not perceive this as foreign accented (Hazan & Boulakia, 1993). This may explain why the L1D–L2G speakers did not suppress prevoicing in L2-German. The finding that the L1G–L2D speakers did not acquire consistent prevoicing in Dutch asks for additional explanations that can be related to articulatory complexity, as discussed in detail in section 5.3.

5.5 Limitations

The present study comes with two limitations. First, the amount and contexts of L2 exposure are confounded with the speakers' L1: as a result of the couples living in the Netherlands, all L1-German bilinguals were exposed more to Dutch than all L1-Dutch bilinguals were exposed to German. Second, the genders were not well balanced across groups: more L1-German bilinguals were female, and more L1-Dutch bilinguals were male. Although all analyses included the variable *Gender*, the uneven distribution of males and females across groups limits statistical power for this variable, as well as for the interactions between *Gender* and *Language* or *Gender* and *Language Background*. These limitations do not affect the main conclusions we can draw from the present study because the relation between the degree of immersion and the degree of nativelikeness is not dependent on whether a bilingual speaks Dutch or German as L1. In addition, we focused on the two bilingual groups individually with respect to both their specific acquisition tasks (acquiring a prevoicing or aspirating L2) and the circumstances of their language learning and use (immersed in the society and the home or exclusively in the home). This allowed us to better understand the way in which each group extended or restructured their phonetic space to accommodate L1 and L2 plosives. As we followed this approach for each group individually, the interpretation is not dependent on the above-mentioned confounding variables. Fully disentangling the effects of the language-learning task and the language-learning circumstances will be a task for future research and would require testing an additional group of Dutch–German couples living in Germany.

6 Conclusions

The present study provided new insight into phonetic differentiation between L1 and L2, as well as L2 attainment and L1 attrition by comparing VOT productions of two groups of L2 speakers who differed in their degree of L2 immersion. Both groups used their L1 and L2 at home, but differed in their L1 vs. L2 use outside the home. Referencing the L2 speakers' speech to L1 speech of their immediate environment, rather than to a monolingual reference group, addressed the question to what extent the L2 speakers had been able to acquire the L2 from the input that is available to them. The results show that both immersion contexts allowed L2 speakers to restructure their phonetic space to accommodate old L1 and new L2 phonetic categories for ‘voiceless’ plosives. Only the L1G–L2D speakers who were frequently exposed to Dutch in a variety of contexts and by multiple speakers in their country of residence restructured their phonetic space to accommodate new L2-Dutch VOT for both ‘voiceless’ and ‘voiced’ plosives. The acquisition of language-specific VOT did not automatically go hand-in-hand with native-like L2 acquisition. Even when the L2 plays a crucial role in everyday life, L1 phonetic attrition seems to be prevented by frequent use of and exposure to the L1 in a variety of contexts and multiple speakers, for example, at the workplace. Combining speech data of bilinguals with L1-Dutch and bilinguals with L1-German for both ‘voiceless’ and ‘voiced’ plosives revealed that success in acquiring native-like VOT in L2 and maintaining native-like VOT in L1 may be limited to VOT in the short lag range.

References

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning, 59*, 249–306.

Abramson, A. S., & Lisker, L. (1973). Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics, 1*, 1–8.

Allen, G. D. (1985). How the young French child avoids the pre-voicing problem for wordinitial voiced stops. *Journal of Child Language*, *12*, 37–46.

Bergmann, C., Nota, A., Sprenger, S. A., & Schmid, M. S. (2016). L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics, 58*, 71–86.

Boersma, P., & Weenink, D. (2015). *Praat: Doing phonetics by computer. Version 5.4.04*. Retrieved from http://www.praat.org/

Bongaerts, T., van Summeren, C., Planken, P., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition, 19*, 447–465.

Bortolini, U., Zmarich, C., Fior, R., & Bonifacio, S. (1995). Word-initial voicing in the productions of stops in normal and preterm Italian infants. *International Journal of Pediatric Otorhinolaryngology*, *31*, 191–206.

Brennan, E., Ryan, E., & Dawson, W. (1975). Scaling of apparent accentedness by magnitude estimation and sensory modality matching. *Journal of Psycholinguistic Research*, *4*, 27–36.

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics, 40*, 249–268.

Cook, V. (2007). The goals of ELT: Reproducing native-speakers or promoting multicompetence among second language users? In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching, Part 1, Vol 15* (pp. 237–248). New York, NY: Springer.

de Leeuw, E. (2014). Maturational constraints in bilingual speech. In E. M. Thomas & I. Mennen (Eds.), *Advances in the study of bilingualism* (pp. 25–37). Clevedon: Multilingual Matters.

de Leeuw, E., Mennen, I., & Scobbie, J. M. (2012). Singing a different tune in your native language: First language attrition of prosody. *International Journal of Bilingualism, 16*, 101–116.

de Leeuw, E., Mennen, I., & Scobbie, J. M. (2013). Dynamic systems, maturational constraints, and phonetic attrition. *International Journal of Bilingualism*, *17*, 683–700.

5

de Leeuw, E., Schmid, M. S., & Mennen, I. (2010). The effects of contact on native language pronunciation in an L2 migrant setting. *Bilingualism: Language and Cognition, 13*, 33–40.

Ferguson, C., & Garnica, O. (1975). Theories of phonological development. In E. H. Lenneberg & E. Lenneberg (Eds.), *Foundations of language development, Vol. 1* (pp. 153–180). New York, NY: Academic Press.

Fischer-Jørgensen, E. (1976). Some data on North German stops and affricates. *Annual Report of the Institute of Phonetics of the University of Copenhagen, 10*, 149–200.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America,* 106, 1511–1522.

Flege, J. E. (1980). Phonetic approximation in second language acquisition. *Language Learning, 30*, 117–134.

Flege, J. E. (1981). The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly,* 15, 443–445.

Flege, J. E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America, 76*, 692–707.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*, 47–64.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America, 89*, 395–411.

Flege, J. E. (1995). Second language speech learning. Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Flege, J. E., & Eefting, W. (1987a). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics, 15*, 67–83.

Flege, J. E., & Eefting, W. (1987b). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication, 6*, 185–202.

Flege, J. E., Frieda, E. M., Walley, A. C., & Randazza, L. A. (1998). Lexical factors and segmental accuracy in second language speech production. *Studies in Second Language Acquisition, 20*, 155–187.

Flege, J. E., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America, 76*, 707–721.

Flege, J. E., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech, 24*, 125–146.

Freed, B. F. (1982). Language loss: Current thoughts and future directions. In R. D. Lambert & B. F. Freed (Eds.), *The loss of language skills* (pp. 1–5). Rowley: Newbury House.

Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology, 42*, 120–133.

Gollan, T. H., Starr, J., & Ferreira, V. S. (2015). More than use it or lose it: The number of speakers effect on heritage language proficiency. *Psychonomic Bulletin & Review, 22*, 147–155.

Hamann, S., & Seinhorst, K. (2016). Prevoicing in Standard German plosives: Implications for phonological representations? Thirteenth Old World Conference in Phonology (OCP), Budapest, Hungary, 13–16 January.

Hazan, V. L., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech, 36*, 17–38.

Hopp, H., & Schmid, M. S. (2013). Perceived foreign accent in first language attrition and second language acquisition: The impact of age of acquisition and bilingualism. *Applied Psycholinguistics, 3*, 361–394.

Jessen, M. (1998). *Phonetics and phonology of tense and lax obstruents in German*. Amsterdam: Benjamins.

Kager, R., van der Feest, S., Fikkert, P., Kerkhoff, A., & Zamuner, T. (2007). Representations of [voice]: Evidence from acquisition. In J. van de Weijer & E. J. van der Torre (Eds.), *Voicing in Dutch: (De)voicing – phonology, phonetics, and psycholinguistics* (pp. 41–80). Amsterdam: Benjamins.

Keating, P. (1984). Phonetic and phonological representation of stop consonant voicing. *Language, 60,* 286–319.

Kewley-Port, D., & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics, 2*, 194–210.

Khattab, G. (2000). VOT in English and Arabic bilingual and monolingual children. In D. Nelson & P. Foulkes (Eds.), *Leeds Working Papers in Linguistics and Phonetics, 8* (pp. 95–122). Retrieved from http://www.leeds.ac.uk/arts/homepage/387/leeds_working_papers_in_linguistics_and_phonetics

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 12577–12580.

Kohler, K. (1977). *Einführung in die Phonetik des Deutschen*. Berlin: Schmidt.

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition, 9*, 119–135.

Kupisch, T., Akpınar, D., & Stöhr, A. (2013). Gender assignment and gender agreement in adult bilinguals and second language learners of French. *Linguistic Approaches to Bilingualism, 3*, 150–179.

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384–422.

Macken, M. A., & Barton, D. (1980a). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language, 7*, 41–74.

Macken, M. A., & Barton, D. (1980b). The acquisition of the voicing contrast in Spanish: A phonetic and phonological study of word-initial stop consonants. *Journal of Child Language, 7*, 433–458.

MacLeod, A. A. (2016). Phonetic and phonological perspectives on the acquisition of voice onset time by French-speaking children. *Clinical Linguistics & Phonetics, 30*, 584–598.

MacLeod, A. A., & Stoel-Gammon, C. (2009). The use of voice onset time by early bilinguals to distinguish homorganic stops in Canadian English and Canadian French. *Applied Psycholinguistics, 30*, 53–77.

Major, R. C. (1987). English voiceless stop production by speakers of Brazilian Portuguese. *Journal of Phonetics, 15*, 197–202.

Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal, 76*, 190–208.

Mayr, R., Price, S., & Mennen, I. (2012). First language attrition in the speech of Dutch-English bilinguals: The case of monozygotic twin sisters. *Bilingualism: Language and Cognition, 15*, 687–700.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics, 32*, 543–563.

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

Riney, T., & Takagi, Y. (1999). Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning, 49*, 275–302.

Rothman, J. (2007). Heritage speaker competence differences, language change, and input type: Inflected infinitives in heritage Brazilian Portuguese. *International Journal of Bilingualism, 11*, 359–389.

Rothman, J., & Treffers-Daller, J. (2014). A prolegomenon to the construct of the native speaker: Heritage speaker bilinguals are natives too! *Applied Linguistics, 35*, 93–98.

Ryalls, J., Zipprer, A., & Baldauff, P. (1997). A preliminary investigation of the effects of gender and race on voice onset time. *Journal of Speech, Language, and Hearing Research, 40*, 642–645.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, 25*, 421–436.

Schmid, M. S. (2004). First language attrition: The methodology revised. *International Journal of Bilingualism, 8*, 239–255.

Schmid, M. S., Gilbers, S., & Nota, A. (2014). Ultimate attainment in late second language acquisition: Phonetic and grammatical challenges in advanced Dutch-English bilingualism. *Second Language Research, 30*, 129–157.

Schmid, M. S., & Köpke, B. (2007). Bilingualism and attrition. In B. Köpke, M. S. Schmid, M. Keijzer & S. Dostert (Eds), *Language attrition: Theoretical perspectives* (pp. 1–7). Amsterdam: Benjamins.

Schoonmaker-Gates, E. (2015). On voice-onset time as a cue to foreign accent in Spanish: Native and nonnative perceptions. *Hispania, 98*, 779–791.

Scovel, T. (1969). Foreign accent, language acquisition and cerebral dominance. *Language Learning, 19*, 245–254.

Seidl, A., Onishi, K. H., & Cristia, A. (2014). Talker variation aids young infants' phonotactic learning. *Language Learning and Development, 10*, 297–307.

Simon, E. (2009). Acquiring a new second language contrast: An analysis of the English laryngeal system of native speakers of Dutch. *Second Language Research, 25*, 377–408.

Simon, E., & Leuschner, T. (2010). Laryngeal systems in Dutch, English, and German: A contrastive phonological study on second and third language acquisition. *Journal of Germanic Linguistics, 22*, 403–424.

Stock, D. (1971). *Untersuchungen zur Stimmhaftigkeit hochdeutscher Phonem-realisationen*. Hamburg: Buske.

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017). Heritage language exposure impacts voice onset time of Dutch-German simultaneous bilingual preschoolers. *Bilingualism: Language and Cognition.* Advance online publication. doi: 10.1017/S1366728917000116

Sundara, M., Polka, L., & Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition, 9*, 97–114.

Ulbrich, C., & Ordin, M. (2014). Can L2-English influence L1-German? The case of post-vocalic /r/. *Journal of Phonetics*, *45*, 26–42.

van Alphen, P. M., & Smits, R. (2004). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics, 32*, 455–491.

Ventureyra, V. A. G., Pallier, C., & Yoo, H. (2004). The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics, 17*, 79–91.

Yu, V. Y., De Nil, L. F., & Pang, E. W. (2015). Effects of age, sex and syllable number on voice onset time: Evidence from children's aspirated stops. *Language and Speech*, *58*, 152–167.

Appendix. Output of the statistical models

RQ 1: Differentiation of VOT in L1 and L2

Table 7. L1G-L2D speakers.

Voiceless plosives				
	β	*SE*	*t*	*p*
Intercept	44.23	3.31	13.36	< .001
Language	16.22	2.41	6.72	< .001
Gender	2.69	2.48	1.08	> .250
WordLength	2.49	1.32	1.88	.060
PoA_LC	-15.05	3.11	-4.84	< .001
PoA_CD	5.93	3.02	1.97	.049
Language*Gender	-1.65	1.28	-1.29	.197
Language*WordLength	2.25	1.30	1.73	.084
Language*PoA_LC	-4.98	2.90	-1.72	.085
Language*PoA_CD	-6.37	2.87	-2.22	.026
Voiced plosives				
	β	*SE*	*z*	*p*
Intercept	-0.69	0.49	-1.39	.165
Language	0.95	0.34	2.84	.005
Gender	1.41	0.48	2.94	.003
PoA	-0.34	0.16	-2.11	.035
Language*Gender	0.37	0.32	1.15	.250
Language*PoA	-0.21	0.16	-1.32	.187

5

Table 8. L1D-L2G speakers.

Voiceless plosives				
	β	*SE*	*t*	*p*
Intercept	37.46	2.54	14.74	< .001
Language	13.83	2.44	5.68	< .001
Gender	2.58	1.42	1.82	.069
WordLength	2.28	1.26	1.81	.070
PoA_LC	-17.63	3.23	-5.46	< .001
PoA_CD	2.06	3.12	0.66	> .250
Language*Gender	2.09	1.13	1.85	.064
Language*WordLength	2.60	1.25	2.07	.039
Language*PoA_LC	-4.75	2.97	-1.60	.110
Language*PoA_CD	-5.68	2.93	-1.94	.052
Voiced plosives				
	β	*SE*	*z*	*p*
Intercept	-2.76	0.49	-5.59	< .001
Language	0.43	0.28	1.54	.124
PoA	-1.12	0.29	-3.86	< .001
Language*Gender	-0.16	0.26	-0.63	> .250
Language*PoA	-0.25	0.18	-1.36	.174

RQ 2: L2 VOT attainment

Table 9. L1G-L2D speakers.

Voiceless plosives				
	β	*SE*	*t*	*p*
Intercept	25.91	2.40	10.80	< .001
LanguageBackground	-2.10	1.45	-1.45	.147
Gender	2.27	1.30	1.74	.082
WordLength	-0.40	1.00	-0.40	> .250
PoA_LC	-11.40	2.64	-4.32	< .001
PoA_CD	9.92	2.66	3.73	< .001
LanguageBackground*Gender	-2.23	1.46	-1.53	.126
LanguageBackground*WordLength	Missing due to convergence problems			
LanguageBackground*PoA_LC	-1.44	1.06	-1.36	.174
LanguageBackground*PoA_CD	-2.30	1.13	-2.04	.041
Voiced plosives				
	β	*SE*	*z*	*p*
Intercept	-2.42	0.49	-4.93	< .001
LanguageBackground	-0.95	0.46	-2.06	.039
Gender	0.38	0.38	1.00	> .250
PoA	-0.68	0.28	-2.43	.015
LanguageBackground*Gender	-0.34	0.38	-0.89	> .250
LanguageBackground*PoA	-0.42	0.22	-1.87	.062

5

Table 10. L1D-L2G speakers.

Voiceless plosives				
	β	*SE*	*t*	*p*
Intercept	54.44	4.19	12.98	< .001
LanguageBackground	-6.57	1.65	-3.97	< .001
Gender	2.14	1.62	1.32	.187
WordLength	5.27	1.76	3.00	.003
PoA_LC	-20.11	5.13	-3.92	< .001
PoA_CD	-0.49	4.88	-0.10	> .250
LanguageBackground*Gender	0.95	2.27	0.40	> .250
LanguageBackground*WordLength	0.13	0.52	0.25	> .250
LanguageBackground*PoA_LC	0.19	1.17	0.16	> .250
LanguageBackground*PoA_CD	Missing due to convergence problems			
Voiced plosives				
	β	*SE*	*z*	*p*
Intercept	-0.87	0.37	-2.34	.019
LanguageBackground	-1.06	0.28	-3.79	< .001
Gender	0.68	0.28	2.44	.015
PoA	-0.66	0.13	-4.93	< .001
LanguageBackground*Gender	-0.92	0.37	-2.49	.013
LanguageBackground*PoA	-0.20	0.13	-1.46	.144

RQ 3: L1 VOT maintenance

Table 11. L1G-L2D speakers.

Voiceless plosives				
	β	*SE*	*t*	*p*
Intercept	67.23	4.69	14.34	< .001
LanguageBackground	-6.94	3.10	-2.24	.025
Gender	-1.25	2.89	-0.43	> .250
WordLength	4.18	1.74	2.41	.016
PoA_LC	-21.52	5.20	-4.14	< .001
PoA_CD	0.75	5.13	0.15	> .250
LanguageBackground*Gender	2.06	2.89	0.71	> .250
LanguageBackground*WordLength	-0.28	0.58	-0.49	> .250
LanguageBackground*PoA_LC	1.27	1.68	0.75	> .250
LanguageBackground*PoA_CD	-1.20	1.52	-0.79	> .250
Voiced plosives				
	β	*SE*	*z*	*p*
Intercept	0.46	0.51	0.90	> .250
LanguageBackground	-0.13	0.50	-0.25	> .250
Gender	1.37	0.50	2.72	.007
PoA	-0.32	0.13	2.53	.011
LanguageBackground*Gender	0.14	0.50	0.28	> .250
LanguageBackground*PoA	-0.12	0.12	-0.94	> .250

5

Table 12. L1D-L2G speakers.

Voiceless plosives				
	β	*SE*	*t*	*p*
Intercept	21.37	1.65	12.93	< .001
LanguageBackground	1.86	1.16	1.60	.110
Gender	0.59	0.62	0.95	> .250
WordLength	-0.27	0.80	-0.33	> .250
PoA_LC	-12.95	1.93	-6.73	< .001
PoA_CD	7.59	1.95	3.90	< .001
Language*Gender	-1.16	0.62	-1.87	.061
Language*WordLength	-0.02	0.32	-0.07	> .250
Language*PoA_LC	-0.12	1.00	-0.12	> .250
Language*PoA_CD	0.18	1.07	0.17	> .250
Voiced plosives				
	β	*SE*	*z*	*p*
Intercept	-3.16	0.45	-7.02	< .001
LanguageBackground	-0.06	0.44	-0.13	> .250
Gender	0.37	0.37	1.00	> .250
PoA	-0.94	0.25	-3.81	< .001
LanguageBackground*Gender	-0.37	0.37	-1.01	> .250
LanguageBackground*PoA	-0.17	0.23	-0.74	> .250

Chapter 6

Bilingual preschoolers’ speech is associated with non-native maternal language input

Based on:

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (under review). *Bilingual preschoolers’ speech is associated with non-native maternal language input.*

Abstract

Bilingual children are often exposed to non-native speech through their parents. Yet, little is known about the relation between bilingual preschoolers’ speech production and their speech input. The present study investigated productions of voice onset time (VOT) by Dutch-German bilingual preschoolers and their late bilingual mothers. The findings reveal an association between maternal VOT and bilingual children’s VOT in the heritage language German as well as in the majority language Dutch. By contrast, no input–production association was observed in VOT productions of monolingual German and monolingual Dutch children. The results of this study provide the first empirical evidence that non-native and attrited maternal speech contributes to the often-observed linguistic differences between bilingual children and their monolingual peers.

6

1 Introduction

The considerable amount of language input that children receive from their parents is an important factor in children's language learning (Hart & Risley, 1995; see Snow, 2014, for an overview). In particular, maternal input appears to be crucial in young children's language development. For example, the amount of maternal language input is positively correlated with monolingual children's lexical knowledge (Hoff, 2003; Hurtado, Marchman & Fernald, 2008; Rowe, 2008, 2012; Rowe, Raudenbush & Goldin-Meadow, 2012). Despite evidence that the speech of children who are exposed to variable sociophonetic input reflects sociophonetic details similar to their mothers' speech (Foulkes, Docherty & Watt, 1999), the direct influence of maternal input on children's linguistic development beyond the lexicon has received little attention to this day.

While the mother may be one of the most important among several input providers for a monolingual child, she may be the only input provider for a bilingual child in one language. Such contexts may arise, for example, when a child is born to binational parents, and acquires the mother's native language as a heritage language.

Crucially, children raised bilingually by binational parents are commonly exposed to non-native language input from their parents in both the majority language and the heritage language: the parents are likely to speak each other's native language (L1) as second language (L2), but with non-native phonetic characteristics (e.g., Flege, 1987, 1991; Flege & Eefting, 1987; Flege & Port, 1981; Simon & Leuschner, 2010). Depending on their L2 use and the length of residence outside of their L1 community, the parents of bilingual children may also speak their L1 differently than monolingual parents (Bergmann, Nota, Sprenger & Schmid, 2016; Chang, 2012; de Leeuw, Mennen & Scobbie, 2012; Flege, 1987; Flege & Hillenbrand, 1984; Major, 1992; Mayr, Price & Mennen, 2012; Mennen, 2004; Sancier & Fowler, 1997; Ulbrich & Ordin, 2014; Ventureyra, Pallier & Yoo, 2004). Changes to the L1, for example because the L2 widely replaces the use of the L1 after emigration, are known as *first language attrition* (Freed, 1982; Schmid, 2004).

The present study focuses on a largely understudied aspect of bilingual first language acquisition by testing whether differential phonetic aspects in the speech of late bilingual mothers are associated with their children's speech production. The term *differential* refers to divergences in bilingual speech from the monolingual norm (Kupisch & Rothman, 2016). We specifically focus on productions of voice onset time (VOT) of mothers and children speaking Dutch and German in the Netherlands.

Bilingual children often produce VOT differently from their monolingual peers when they acquire two languages that require different VOT values, such as Dutch and German (Deuchar & Clark, 1996; Fabiano-Smith & Bunta, 2012; Johnson & Wilson, 2002; Kehoe, Lleó & Rakow, 2004; Khattab, 2000, 2003). VOT is a cue to plosive voicing and describes the time between the release of a plosive's closure and the onset of vocal fold vibration (Lisker & Abramson, 1964). 'Voiceless' plosives (/p/, /t/, /k/) and 'voiced' plosives (/b/, /d/, /g/) primarily differ by means of VOT. For example, shortening of the VOT of the /p/ in the German word /piːɐ̯/ <Pier> (pier) can cause it to be perceived as a /b/, which also changes the word's meaning to /biːɐ̯/ <Bier> (beer). The phonetic realization of 'voiceless' and 'voiced' plosives can differ between languages. For example, German and English 'voiceless' plosives have *aspiration,* which means that they are produced with long VOT values. German and English 'voiced' plosives are realized with short positive VOT or *short lag* VOT. Occasionally, German and English 'voiced' plosives are produced with negative VOT or *prevoicing*, which means that the vocal folds start vibrating prior to the release of the plosive. By contrast, 'voiceless' plosives in Dutch, Arabic, Spanish and Italian typically have short lag VOT. 'Voiced' plosives in these languages require prevoicing, but it is nevertheless common that adult native speakers of a prevoicing language fail to produce prevoicing for a small percentage of their 'voiced' plosives (Khattab, 2003; van Alphen & Smits, 2004).

When the speech of bilingual children differs from the speech of monolingual children, this has often been interpreted as resulting from cross-linguistic influence (CLI), that is, the interplay between two languages during language processing (Fabiano & Goldstein, 2005; Fabiano-Smith & Bunta, 2012; Kehoe, 2002; Kehoe et al., 2004; Lleó & Kehoe, 2002; Müller & Hulk, 2001; Paradis & Genesee, 1996). However, rather than emerging exclusively from CLI between the child's two languages, deviances of bilinguals' speech from the monolingual reference point may in part be related to the non-native and attrited language input the bilingual child receives in each of her languages (Fish, García-Sierra, Ramírez-Esparza & Kuhl, 2017).

Similarities between the phonetic properties of bilingual children's input and their own speech production were previously observed in four case studies of bilingual and trilingual children whose language input in one language was limited to a small number of speakers (Deuchar & Clark, 1996; Khattab, 2003; Klinger, 1962; Mayr & Montanari, 2015). Klinger (1962) describes the influence of atypical speech input on a bilingual child's global accent. John, a Spanish-English bilingual child, acquired English exclusively from his older cousin, whose speech was atypical due to a cleft palate. While John followed typical language development in Spanish, he adopted his cousin's cleft palate speech symptoms when he spoke English. Following a speech therapist's advice, John was no longer exposed to his cousin's speech from the age of 3;6 (years; months), and instead heard English from other speakers. At age 5;0, John's English speech was no longer characterized by cleft palate speech symptoms. This report illustrates the link between children's language input and their speech productions, and also highlights the impact of language input that comes from a single source, rather than from a diverse set of speakers.

Similarities between maternal input and a bilingual child's speech production based on acoustic measures of VOT were first reported by Deuchar and Clark (1996). During a period of six months, they recorded the speech of the parents and their English-Spanish bilingual child raised in England. By the end of data collection when the child was aged 2;3 years, she had developed a voicing contrast between short lag VOT and aspiration in English. In Spanish, a covert voicing contrast within the short lag-range appeared to be developing. The mother, a native speaker of English who spoke Spanish as L2, made a similar covert voicing contrast in the short lag VOT range when she spoke Spanish. Given that target-like prevoicing is generally acquired after the age of 2;3 years in monolingual acquisition, the authors did not draw strong conclusions about a possible relation between maternal input and the child's VOT production.

Input effects of maternal VOT on bilingual children's VOT productions of two English-Arabic bilingual children aged seven and ten years were discussed by Khattab (2003). The children acquired Arabic as a heritage language in the United Kingdom from their parents, and one child resorted to a differential phonetic realization of the voicing contrast in Arabic. Instead of producing 'voiced' plosives with prevoicing, the child produced either prenasalized plosives or implosives, for which the airstream flows inward the mouth. Acoustic analyses of the mother's speech revealed that she often produced prenasalization or implosives instead of prevoiced plosives as well. As these realizations diverge from monolingual acquisition patterns, Khattab's study appears to be the first that demonstrates a link between specific acoustic patterns in bilingual children's speech and their maternal language input.

A nuanced view on input characteristics comes from a study on two trilingual sisters' VOT productions in all three languages (Mayr & Montanari, 2015). The speech of the English-Italian-Spanish trilingual sisters and their primary input providers had been recorded when the children were aged 6;8 and 8;1. Exposure to

English came from the father, and the broad social environment. Moreover, the sisters were exposed to English at their Italian immersion school, in which English and Italian were the languages of instruction. Italian input came from their Italian mother and from their Italian immersion school. Spanish input was exclusively provided by their monolingual Spanish-speaking nanny. The children's English VOT was target-like and broadly in line with their father's productions. Their Spanish VOT productions were also target-like and very closely matched those of the nanny. Non-target like productions were only observed in Italian, in which the children produced 'voiceless' plosives with longer than target-like VOT, which can be regarded as more English-like VOT. It is intriguing that CLI seems to be present from English to Italian, while the children's weakest language – Spanish – appeared to be unaffected by CLI from the majority language English. The authors suggested that the children's differential Italian VOT productions are related to their exposure to non-native Italian at their Italian immersion school, which was predominantly attended by children for whom Italian was the L2. In Spanish, the sisters were only exposed to a single speaker, who provided monolingual-like input, and the children therefore acquired monolingual-like Spanish VOT themselves. The study of Mayr and Montanari suggests that the input contributes to children's monolingual-like or differential speech production in a multilingual acquisition setting.

Taken together, the four studies observed striking similarities between the language input and a bilingual child's speech production. Given the small sample sizes, these studies were descriptive and did not allow for statistical analyses of association between the input and the children's speech. Such analyses necessarily require a larger sample of children and are essential to provide evidence for the claim that phonetic characteristics of the speech input are indeed associated with a child's speech production.

In sum, non-native language input appears to be a crucial – but largely unexplored – factor that should be acknowledged when comparing the linguistic skills of bilingual and monolingual children. The present study is the first to address whether non-native and attrited maternal speech input is associated with bilingual preschoolers' speech production.

1.1 The current study

The current study investigates whether individual variation in VOT production of Dutch-German simultaneous bilingual children reflects variation in the VOT production of their late bilingual mothers who speak German as L1 and Dutch as

L2. All participants lived in the Netherlands and the children were therefore immersed in a Dutch language environment. They were exposed to German only in their family, and primarily through their German mother.

Bilingual children often produce VOT differently from monolingual children in at least one of their languages (Deuchar & Clark, 1996; Fabiano-Smith & Bunta, 2012; Johnson & Wilson, 2002; Kehoe et al., 2004; Khattab, 2000, 2003). Most of the bilingual children examined in the above-mentioned studies were raised in one-parent-one-language families, and in order to communicate with each other, at least one of their parents spoke the other parent's native language as L2. Adult L2 VOT typically diverges from monolinguals' VOT (Flege, 1987, 1991; Flege & Eefting, 1987; Flege & Port, 1981), and non-native VOT productions have recently been observed in the speech directed toward bilingual infants (Fish et al., 2017). Furthermore, the parents' VOT may even diverge from monolingual native speakers' VOT in their native language due to language attrition, presumably resulting from reduced language contact with other native speakers (Flege, 1987; Major, 1992; Mayr et al. 2012; Sancier & Fowler, 1997). Consequently, bilingual children are likely exposed to non-native and attrited VOT from their parents.

The present study builds on two recent studies that investigated VOT productions of bilingual children (Stoehr, Benders, van Hell & Fikkert, 2017b / Chapter 2) and their bilingual parents (Stoehr, Benders, van Hell & Fikkert, 2017a / Chapter 5). These two studies, which are summarized below, revealed conspicuous similarities at the group-level between the VOT productions of bilingual children and the non-native and attrited VOT productions of their mothers.

In a large-scale study, analyses of the speech of Dutch-German simultaneous bilingual preschoolers aged 3;6 to 6;0 and aged-matched Dutch and German monolingual children revealed that the VOT of bilinguals differed from monolinguals' VOT in both languages (Stoehr et al., 2017b / Chapter 2). In the heritage language German, the bilingual children, compared to monolingual children, produced more 'voiced' plosives with prevoicing, and they also produced 'voiceless' plosives with shorter VOT. In the majority language Dutch, the differences between bilinguals and monolinguals were less pronounced. The bilinguals produced 'voiced' plosives with prevoicing less consistently than monolingual Dutch children. However, even the monolingual Dutch children did not yet reach adult-like consistency in prevoicing for 'voiced' plosives, which is in line with previous research reporting late mastery of prevoicing in monolingual acquisition (Khattab, 2000; MacLeod, 2016). The bilingual children produced Dutch 'voiceless' plosives only with slightly longer VOT than monolingual children, and these productions still fell within monolingual-like ranges. In sum, the observed differences in VOT production between bilingual and monolingual children in both German and Dutch can be interpreted as resulting from CLI that operates from Dutch to German, and to a lesser extent also from German to Dutch.

A factor contributing to the bilingual children's differential VOT productions may be their exposure to non-native and attrited speech. The bilingual children's mothers who spoke German as L1 and Dutch as L2 exhibited a VOT production pattern that was strikingly similar to the children's VOT production at the group-level (Stoehr et al., 2017a / Chapter 5)[17]. The mothers had been living in the Netherlands for several years at the time of testing and it appeared that they were affected by phonetic attrition of L1 VOT: they produced German 'voiceless' plosives with shorter, and therefore more Dutch-like VOT compared to monolingual speakers of German. 'Voiced' plosives, on the other hand, were produced with variable VOT including prevoicing (about one third of all productions) and short lag VOT by the bilingual mothers and monolingual German-speaking women alike.
Moreover, the bilingual mothers also produced non-native VOT in their L2 Dutch[18]. They produced longer VOT values in 'voiceless' plosives and produced fewer 'voiced' plosives with prevoicing in comparison to monolingual women.

Given these deviations from monolingual-like VOT productions in both the bilingual children and their mothers, the present study investigates whether bilingual children's differential VOT production is associated with their own mothers' non-native or attrited VOT production. To assess whether such an association is limited to a bilingual acquisition context or arises during language acquisition in general, we ask whether an association between VOT production and language input can be observed in monolingual child-mother dyads.

Regarding the outcome of this study, we hypothesize that maternal VOT is associated with the VOT productions of both bilingual and monolingual children. This hypothesis is based on previous research, which reported maternal input effects in monolingual children's lexical growth (Hoff, 2003; Hurtado et al., 2008; Rowe, 2008, 2012; Rowe et al., 2012) and observations of similarities between phonetic

[17] The participants investigated in Stoehr et al. (2017a / Chapter 5) furthermore included six fathers who spoke L1 German and L2 Dutch. For the purpose of the present study, we verified that the same pattern of results holds after exclusion of those additional six participants.

[18] In Stoehr et al. (2017a / Chapter 5), the parents' L2-Dutch VOT was compared to the VOT of Dutch native-speakers who spoke German as L2 to address their acquisition outcomes. Given that the present study is concerned with differences in the input of bilingual and monolingual children, the data of Stoehr et al. (2017a / Chapter 5) have been reanalyzed using monolingual native speakers as reference group. These findings are reported in the present chapter.

aspects of the input and the speech of monolingual children (Foulkes et al., 1999) and bilingual children (Deuchar & Clark, 1996; Khattab, 2003; Klinger, 1962; Mayr & Montanari, 2015).

2 Method

2.1 Participants

The participants were a subset of those reported in Stoehr et al., (2017b) / Chapter 2 and Stoehr et al. (2017a) / Chapter 5. Seventy-four children aged between 3;6 and 6;0 participated in this study: 23 Dutch-German simultaneous bilingual children (12 female, M_{age}=4;8, SD_{age}=9 months), 26 Dutch monolingual children (13 female, M_{age}=4;10, SD_{age}=10 months) and 25 German monolingual children (16 female, M_{age}=4;7, SD_{age}=10 months). In addition, each child's mother participated.

Based on parental report, all children were typically developing and had no speech impairments or delays, and no neurological, auditory or cognitive impairments. All bilingual children lived in the Netherlands from birth. The mothers of all bilingual children were native speakers of German. Twenty bilingual children had a Dutch father, and the remaining three bilingual children had a German father. Out of the three bilingual children whose parents were both native speakers of German, two were exposed to Dutch from native speakers from birth. The third child started being exposed to Dutch frequently through native speakers at 0;6 years. On average, the bilingual children were exposed to German for 45% of the day (range 11% – 69%, *SD*=15%) at the time of testing, as determined by the Bilingual Language Experience Calculator (BiLEC; Unsworth, 2013).

An additional 23 children (10 bilinguals and 13 monolinguals) had been tested, but were excluded from the analysis. The bilinguals were excluded either due to exposure to a third language (*N*=3) or onset of bilingualism after the first year of life (*N*=1). In addition, bilingual children born to a Dutch-speaking mother were also excluded (*N*=6) to obtain a more homogeneous data set for this study. The 13 monolingual children were excluded either due to inability to complete the task (*N*=1), experimenter error (*N*=1), exposure to non-native speakers at home (*N*=4), or missing speech data of the mother (*N*=7).

The mothers of the bilingual children learned Dutch at an average age of 23 years (range 8 – 33 years, *SD*=6 years) when they moved to the Netherlands. Twenty-one mothers reported frequent use of German and Dutch. One mother reported occasional use of Dutch and one mother did not provide information on her usage of Dutch. The mothers rated themselves as very proficient in speaking Dutch (on a scale from 0 [virtually no fluency] to 5 [native fluency]: *M*=4.0, *SD*=0.64, range 3 – 5), and almost native-like in understanding Dutch (on a scale from 0 [almost no understanding] to 5 [native understanding]: *M*=4.4, *SD*=0.51, range 4 – 5). The bilingual families lived within a radius of 100 kilometers from Nijmegen in the Central Eastern Netherlands and were tested at their homes.

Four of the monolingual German mothers knew some Dutch, but none of them reported regular use of a language different from German. The German monolingual child-mother dyads were tested in Central Western Germany (*N*=23) and Northern Germany (*N*=2).

Two monolingual Dutch mothers reported speaking some German, and three reported speaking English sporadically. All Dutch monolingual child-mother dyads were tested in Nijmegen or its periphery. Although most mothers in both groups did not report speaking additional languages, schooling in Germany and the Netherlands requires English language classes in high school, suggesting that all mothers knew some English.

2.2 Materials and procedure

The children's speech elicitation is described in Chapter 2, section 2.2, and the elicitation of the mother's speech is described in Chapter 5, section 2.2.

2.3 Recordings and VOT measurements

Recordings and VOT measurements are described in Chapter 2, section 2.3 and Chapter 5, section 2.3.

2.4 Statistical analyses

Mixed effects linear and logistic regression analyses were performed in *R* (R Core Team, 2013). Separate models were run for bilingual and monolingual children. An α-level of .05 was adopted throughout. Per group, three analyses were conducted: the main model included data of 'voiceless' and 'voiced' plosives. However, VOT of 'voiced' plosives was bimodally distributed, which violates the assumption of normality in linear regression. For this reason, data were split by voicing. The sub-model on 'voiceless' plosives was designed as a mixed effects linear regression model in line with the main model. The sub-model on 'voiced' plosives was designed as a mixed effects logistic regression model accounting for the bimodal data distribution. The main model and the two sub-models are described below.

Main models. In the main mixed effects linear regression models, the dependent variable was the children's VOT for each target word they produced in the study. As fixed effects, the bilingual model used *Maternal VOT* (continuous, in ms), *Voicing* (voiced=-1, voiceless=1), *Language* (Dutch=-1, German=1), *Exposure to German* (continuous, in percent; inversely related to Exposure to Dutch), *Place of Articulation* (labial=1 vs. coronal=0; and coronal=0 vs. dorsal=1) and *Word Length* (monosyllabic=-1, disyllabic=1). Based on the results of Stoehr et al. (2017a / Chapter 5, b / Chapter 2), the following interaction terms were included: a four-way interaction term between *Maternal VOT* * *Voicing* * *Language* * *Exposure to German*, a three-way interaction between *Language* * *Voicing* * *Place of Articulation* (labial vs. coronal), a two-way interaction between *Language* * *Place of Articulation* (coronal vs. dorsal), and a three-way interaction between *Language* * *Voicing* * *Word Length*. As random effects, the model included intercepts for *Child* and *Target Word*, as well as by-*Child* random slopes for *Language* and *Voicing*. The model for monolingual children was the same with the exception that it excluded the fixed effect *Exposure to German* and also the random slopes for *Language*.

Sub-models 'voiceless' plosives. The sub-models on 'voiceless' plosives were based on the main model, but excluded the fixed effect *Voicing* and the by-*Child* random slopes for *Voicing*. Moreover, *Voicing* was removed from the interaction terms that were included in the main model.

Sub-models on 'voiced' plosives. The sub-model on 'voiced' plosives was a mixed effects logistic regression accounting for the bimodal distribution of VOT in 'voiced' plosives. In this analysis, the dependent variable was the proportion of 'voiced' plosives the children produced with prevoicing. As fixed effects, the bilingual model included *Maternal VOT* (prevoiced=0, devoiced=1), *Language* (Dutch=-1, German=1), *Exposure to German* (in percent), and *Place of Articulation* (labial=1, coronal=0). Based on the results of Stoehr et al. (2017a / Chapter 5, b / Chapter 2), a three-way interaction term between *Maternal VOT* * *Language* * *Exposure to German* and a two-way interaction term between *Language* * *Place of Articulation* were entered. As random effects, the model included intercepts for *Child* and *Target Word* and by-*Child* random slopes for *Language*. The model for monolingual children was the same with the exception that it excluded the fixed effect *Exposure to German* and also the random slopes for *Language*.

In the following, the results regarding the main effect of *Maternal VOT* and interactions involving *Maternal VOT* are reported. Other significant main effects and interactions regarding the other independent variables included in the analyses are presented in the Appendix.

3 Results

3.1 Bilinguals

Main model. A mixed effects linear regression analysis based on 2974 observations was conducted to test if maternal VOT productions are associated with the bilingual children's VOT production in 'voiceless' and 'voiced' plosives. The results revealed an association between maternal VOT and the bilingual child's own VOT production (β=0.124, *SE*=0.045, *t*=2.513, *p*=.012). There was no significant interaction between *Maternal VOT* and *Language*, suggesting that the association is present in both Dutch and German (β=0.042, *SE*=0.042, *t*=1.000, *p*>.250).

Sub-model 'voiceless' plosives. A mixed effects linear regression analysis based on 1843 observations was conducted to test if maternal VOT productions are associated with bilingual children's VOT production in 'voiceless' plosives only. This sub-model confirmed the association between maternal VOT and the bilingual child's own VOT production in 'voiceless' plosives (β=0.132, *SE*=0.054, *t*=2.436, *p*=.015). No significant interaction between *Maternal VOT* and *Language* was detected, suggesting that the association is present in both Dutch and German 'voiceless' plosives (β=0.005, *SE*=0.053, *t*=0.100, *p*>.250).

Sub-model 'voiced' plosives. A mixed effects logistic regression analysis based on 1131 observations was conducted to confirm that the effect of maternal VOT holds in a model that accounts for the bimodal distribution of VOT in 'voiced' plosives. However, the model did not detect a significant association between maternal VOT and the children's VOT production for 'voiced' plosives (β=0.205, *SE*=0.231, *z*=0.889, *p*>.250). Moreover, no significant interaction between *Maternal VOT* and *Language* was detected (β=-0.392, *SE*=0.228, *z*=-1.718, *p*=.086).

In sum, the results suggest that the VOT productions of the mothers who speak German as L1 and Dutch as L2 are associated with their children's VOT productions in both Dutch and German. Furthermore, this effect is only detected in bilingual children's productions of 'voiceless' plosives, and it could not be detected in a model which accounts for the bimodal distribution of VOT in 'voiced' plosives.

3.2 Monolinguals

Main model. A mixed effects linear regression analysis based on 3172 observations was conducted to test if maternal VOT is associated with monolingual children's VOT. The model did not detect a significant association between maternal VOT and the monolingual child's own VOT production (β=0.052, *SE*=0.046, *t*=1.149, *p*<.250), but a significant interaction between *Maternal VOT* and *Voicing* was detected (β=0.099, *SE*=0.046, *t*=2.184, *p*=.029). This interaction was investigated by analyzing the data for 'voiceless' and 'voiced' plosives separately. These post hoc analyses did not detect an effect of *Maternal VOT* on monolingual children's VOT productions neither in 'voiceless' plosives (β=0.092, *SE*=0.062, *t*=1.481, *p*=.139) nor in 'voiced' plosives (β=-0.046, *SE*=0.033, *t*=-1.372, *p*=.170). Instead, the interaction seems to reflect differences in the direction of the association between maternal VOT and monolingual children's VOT for 'voiceless' plosives (positive β) and 'voiced' plosives (negative β), neither of which is significant in its own right. In addition, no significant interaction between *Maternal VOT* and *Language* was detected (β=-0.061, *SE*=0.045, *t*=-1.333, *p*=.183).

Sub-model 'voiceless' plosives. A mixed effects linear regression analysis based on 1985 observations was used to test if maternal VOT productions are associated with monolingual children's VOT productions in 'voiceless' plosives only. As reported in the post hoc test accompanying the main model above, no association was detected between maternal VOT and the monolingual child's own VOT productions in 'voiceless' plosives. Moreover, no significant interaction between *Maternal VOT* and *Language* was detected (β=-0.116, *SE*=0.062, *z*=-1.853, *p*=.064).

Sub-model 'voiced' plosives. To account for the bimodal distribution of VOT in the 'voiced' plosives, a mixed effects logistic regression analysis based on 1187 observations was additionally conducted. This model tested if an effect of maternal VOT productions on monolingual children's VOT production is detected in 'voiced' plosives using a model that accounts for the bimodal distribution in the data. Like the main model, the logistic regression model did not detect a significant association between maternal VOT and the monolingual children's VOT production for 'voiced' plosives (β=-0.468, *SE*=0.288, *z*=-1.627, *p*=.104). No significant interaction between *Maternal VOT* and *Language* was detected (β=-0.060, *SE*=0.288, *z*=-0.208, *p*>.250).

4 Discussion

The primary aim of the present study was to test whether individual variation in Dutch-German bilingual preschoolers' VOT production is related to individual variation of VOT in their mothers' native language (L1) and second language (L2) speech. The bilingual children acquired German as a heritage language predominantly from their mothers, who were L1-speakers of German. Dutch was the bilingual children's majority language, and was spoken by their fathers, in the broad social environment, as well as by the bilingual children's mothers who are L2–speakers of Dutch. In addition, this study sought to answer whether such an input–production association arises during language acquisition in general, and is thus also present between the speech of monolingual children and their monolingual mothers. The results of this study represent the first statistical evidence that differential speech input contributes significantly to a bilingual child's differential speech production.

We hypothesized that maternal VOT is associated with the VOT production of bilingual and monolingual children, and this hypothesis was partially confirmed. An association between maternal input and the bilingual children's production was present in the heritage language German, in which the mother was the primary input provider. Moreover, an input–production association was also observed in the bilinguals' majority language Dutch, in which the input was provided by many speakers besides the mother. The input–production associations in German and Dutch, were, however, only detected in productions of 'voiceless' plosives. Against our hypothesis, no input–production association was detected in the speech of monolingual children. In this section, we first discuss reasons that may contribute to the presence of an input–production association in bilingual acquisition, and to the apparent absence of this association in monolingual acquisition. We then discuss why maternal VOT was only detectably associated with the bilingual children's VOT productions of 'voiceless' plosives.

The majority of the bilingual children in this study were raised by parents who speak different native languages. The mothers of the bilingual children in the present study all moved from Germany to the Netherlands before their children were born and used their L2 – Dutch – as means of communication in everyday life. Yet, the mothers produced non-native VOT in the L2 (Stoehr et al., 2017a / Chapter 5). Moreover, the mothers' restricted use of their L1 – German – caused their German VOT for 'voiceless' plosives to become more Dutch-like (Stoehr et al., 2017a / Chapter 5).

The observed association between maternal VOT and the bilingual children's VOT production suggests that bilingual children are affected by their mothers' non-native speech in the majority language Dutch as well as by their mothers' attrited speech in the heritage language German. This finding represents important evidence against the implicit assumption that bilingual children are exposed to similar input as their monolingual peers. The direct influence of phonetic aspects in maternal speech on bilingual children's speech production puts into perspective the often-observed differences in VOT production of bilingual and monolingual children (Deuchar & Clark, 1996; Fabiano-Smith & Bunta, 2012; Johnson & Wilson, 2002; Kehoe et al., 2004; Khattab, 2000, 2003). The results of the present study specifically show that there are factors beyond CLI that can cause the speech of bilingual children to diverge from the speech of monolingual children. The present study furthermore supports previous case studies describing striking similarities between the input and bilingual children's speech production with statistical evidence (Deuchar & Clark, 1996; Khattab, 2003; Klinger, 1962; Mayr & Montanari, 2015). In sum, when phonetic aspects of the speech of bilingual children differ from monolingual children, this can in part be attributed to differences in their input.

One crucial difference between monolingual and bilingual children in this study is the bilingual children's exposure to their mothers' non-native Dutch at home. By contrast, none of the monolingual children were exposed to non-native speakers of their native language in their immediate social environment, as confirmed by parental report. An association between maternal VOT and children's VOT is possibly present in bilingual and monolingual children alike, but there may not be sufficient individual variation in VOT productions of monolingual mothers and children to detect such an association. Previous research showed that maternal input is associated with monolingual children's lexical growth, and this finding supports the hypothesis that an input–production association remained undetected in the present study (Hoff, 2003; Hurtado et al., 2008; Rowe, 2008, 2012; Rowe et al., 2012). Future research is required to test whether input–production associations are detected in a monolingual acquisition context that is likely to yield individual variation, for example involving dialectal or sociophonetic variation.

It appeared in the present data that only the bilingual children's productions of 'voiceless' plosives in both German and Dutch were affected by maternal VOT production. This finding raises the question why bilingual children seem to adopt the VOT of their mothers for 'voiceless' plosives, but not 'voiced' plosives. One explanation is that maternal VOT of 'voiceless' plosives is a target that children can easily match, as short lag VOT and aspiration are relatively less complex in their articulatory gestures than prevoicing, and therefore are also acquired earlier in production (Kewley-Port & Preston, 1974; Macken & Barton, 1980a, b). For German 'voiced' plosives, alternations between short lag VOT and prevoicing were common in the speech of the bilingual mothers and their children, and at the group-level, both mothers and children produced approximately one third of all German 'voiced' plosives with prevoicing. Given that the use of prevoicing in German appears to be idiosyncratic (see Khattab, 2003, for similar thoughts on prevoicing in English), the usage of prevoicing in word-initial singleton plosives does not seem to follow predictable rules in German. These variable input patterns render it unlikely to observe word-specific similarities in mothers' and children's production of prevoicing in German at an individual level.

In Dutch, the bilingual children also produced about one third of all 'voiced' plosives with prevoicing at the group-level, whereas the mothers produced almost two thirds of all Dutch 'voiced' plosives with prevoicing at the group-level. Two factors are likely to hinder direct input–production associations between maternal VOT and child VOT of 'voiced' plosives in Dutch. First, the mothers are non-native speakers of Dutch, and do not produce all of their Dutch 'voiced' plosives with prevoicing. Their use of prevoicing may therefore also be to some extent idiosyncratic in Dutch. As in German, a variable production pattern of 'voiced' plosives with either prevoicing or short lag VOT makes it unlikely to observe a word-specific match in the production of 'voiced' plosives between mothers and their children in Dutch. Second, prevoicing requires complex velopharyngeal adjustments (Kewley-Port & Preston, 1974), which do not appear to be completely mastered by children between 3;6 and 6;0 years of age (Khattab, 2000; MacLeod, 2016). For this reason, the children in this study may motorically not be able to match their mothers' VOT production for 'voiced' plosives in Dutch.

5 Conclusions

The current study provided novel evidence that individual variation in maternal language input is associated with individual variation in bilingual children's speech production. No input–production association was detected in monolingual children's production of VOT, which may result from a lack of variance in monolinguals' VOT production. The association between maternal VOT production and bilingual children's VOT production suggests that linguistic differences between simultaneous bilingual children and their monolingual peers are not exclusively driven by CLI between a bilingual child's linguistic systems. The speech of the bilingual children's mothers, which diverges from monolingual mothers' speech because the bilinguals' mothers are L2-speakers and L1-attriters, contributes to differential speech productions of three-and-a-half to six-year-old bilingual children.

References

Bergmann, C., Nota, A., Sprenger, S. A., & Schmid, M. S. (2016). L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics, 58*, 71–86.

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, *40*, 249–268.

de Leeuw, E., Mennen, I., & Scobbie, J. M. (2012). Singing a different tune in your native language: First language attrition of prosody. *International Journal of Bilingualism, 16*, 101–116.

Deuchar, M., & Clark, A. (1996). Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics, 24*, 351–365.

Fabiano, L., & Goldstein, B. (2005). Phonological transfer in bilingual Spanish-English speaking children. *Journal of Multilingual Communication Disorders, 3*, 56–63.

Fabiano-Smith, L., & Bunta, F. (2012). Voice onset time of voiceless bilabial and velar stops in 3-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics & Phonetics, 26*, 148–163.

Fish, M. S., García-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. (2017). Infant-directed speech in English and Spanish: Assessments of monolingual and bilingual caregiver VOT. *Journal of Phonetics, 63*, 19–34.

Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*, 47–65.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in plosives produced in a second language. *Journal of the Acoustical Society of America, 89*, 395–411.

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics, 15*, 67–83.

Flege, J. E., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America*, *76*, 708–721.

Flege, J. E., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech, 24*, 125–146.

Foulkes, P., Docherty, G., & Watt, D. (1999). Tracking the emergence of structured variation – realisations of (t) by Newcastle children. In P. Foulkes (Ed.), *Leeds Working Papers in Linguistics and Phonetics, 7* (pp. 1–23). Retrieved from

6

http://www.leeds.ac.uk/arts/info/125154/leeds_working_papers_in_linguistics_and_phonetics/1938/volume_7_1999

Freed, B. F. (1982). Language loss: Current thoughts and future directions. In R. D. Lambert & B. F. Freed (Eds.), *The loss of language skills* (pp. 1–5). Rowley: Newbury House.

Grimm, H., & Doil, H. (2000). *ELFRA: Elternfragebögen für die Früherkennung von Risikokindern, Handanweisungen.* Göttingen: Hogrefe.

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Brookes.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*, 1368–1378.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science, 11*, F31–F39.

Johnson, C. E., & Wilson, I. L. (2002). Phonetic evidence for early language differentiation: Research issues and some preliminary data. *International Journal of Bilingualism, 6*, 271–289.

Kehoe, M. (2002). Developing vowel systems as a window to bilingual phonology. *International Journal of Bilingualism, 6*, 315–334.

Kehoe, M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition, 7*, 71–88.

Kewley-Port, D., & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics, 2*, 194–210.

Khattab, G. (2000). VOT in English and Arabic bilingual and monolingual children. In D. Nelson & P. Foulkes (Eds.), *Leeds Working Papers in Linguistics and Phonetics, 8* (pp. 95–122). Retrieved from http://leeds.ac.uk/arts/homepage/387/leeds_working_papers_in_linguistics_and_phonetics

Khattab, G. (2003). Age, input, and language mode factors in the acquisition of VOT by English-Arabic bilingual children. In M. J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the International Congress of Phonetic Sciences, 15* (pp. 3213–3216). Retrieved from https://www.internationalphoneticassociation.org/ icphs-proceedings/ICPhS2003/papers/p15_3213.pdf

Klinger, H. (1962). Imitated English cleft palate speech in a normal Spanish speaking child. *Journal of Speech and Hearing Disorders, 27*, 379–381.

Kupisch, T., & Rothman, J. (2016). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers.

International Journal of Bilingualism. Advance online publication. doi: 10.1177/1367006916654355

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20,* 384–422.

Lleó, C., & Kehoe, M. (2002). On the interaction of phonological systems in child bilingual acquisition. *International Journal of Bilingualism, 6*, 233–237.

Macken, M. A., & Barton, D. (1980a). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language, 7,* 41–74.

Macken, M. A., & Barton, D. (1980b). The acquisition of the voicing contrast in Spanish: A phonetic and phonological study of word-initial stop consonants. *Journal of Child Language, 7,* 433–458.

MacLeod, A. A. (2016). Phonetic and phonological perspectives on the acquisition of voice onset time by French-speaking children. *Clinical Linguistics & Phonetics, 30*, 584–598.

Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal, 76*, 190–208.

Mayr, R., & Montanari, S. (2015). Cross-linguistic interaction in trilingual phonological development: The role of the input in the acquisition of the voicing contrast. *Journal of Child Language, 42,* 1006–1035.

Mayr, R., Price, S., & Mennen, I. (2012). First language attrition in the speech of Dutch-English bilinguals: The case of monozygotic twin sisters. *Bilingualism: Language and Cognition, 15*, 687–700.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics, 32*, 543–563.

Müller, N., & Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition, 4*, 1–21.

Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition*, *18,* 1–25.

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

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language, 35*, 185–205.

6

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development, 83,* 1762–1774.

Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child Development, 83*, 508–525.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, 25*, 421–436.

Schmid, M. S. (2004). First language attrition: The methodology revised. *International Journal of Bilingualism*, *8*, 239–255.

Simon, E., & Leuschner, T. (2010). Laryngeal systems in Dutch, English, and German: A contrastive phonological study on second and third language acquisition. *Journal of Germanic Linguistics, 22*, 403–424.

Snow, C. E. (2014). Input to interaction to instruction: Three key shifts in the history of child language research. *Journal of Child Language*, *41*, 117–123.

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017a). Second language attainment and first language attrition: The case of VOT in immersed Dutch-German late bilinguals. *Second Language Research, 33*, 483–518.

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017b). Heritage language exposure impacts voice onset time of Dutch-German simultaneous bilingual preschoolers. *Bilingualism: Language and Cognition*. Advance online publication. doi: 10.1017/S1366728917000116

Ulbrich, C., & Ordin, M. (2014). Can L2-English influence L1-German? The case of post-vocalic /r/. *Journal of Phonetics*, *45*, 26–42.

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*, 86–110.

van Alphen, P. M., & Smits, R. (2004). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics, 32*, 455–491.

Ventureyra, V. A. G., Pallier, C., & Yoo, H. (2004). The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics, 17*, 79–91.

Appendix. Additional statistical results

Bilinguals

Main model: The main model revealed a significant main effect for *Voicing* (β=26.68, *SE*=7.869, *t*=3.390, *p*<.001), showing that the bilingual children produced longer VOT for 'voiceless' than for 'voiced' plosives. A significant main effect for *Place of Articulation* (labial vs. coronal; β=-14.50, *SE*=3.587, *t*=-4.043, *p*<.001) shows that the bilingual children produced labial plosives with shorter VOT than coronal plosives. A significant three-way interaction between *Language* * *Voicing* * *Exposure to German* (β=0.206, *SE*=0.101, *t*=2.038, *p*=.041) reflects the recent finding that more exposure to German (the heritage language) is associated with longer VOT in German 'voiceless' plosives, while no such effect was observed for German 'voiced' plosives and Dutch 'voiced' and 'voiceless' plosives (Stoehr et al., 2017b / Chapter 2). No other main effects or interactions were significant.

Sub-model 'voiceless' plosives: The model revealed a main effect for *Language* (β=13.46, *SE*=4.507, *t*=2.987, *p*=.003), which shows that the bilingual children produced longer VOT in German 'voiceless' plosives than in Dutch 'voiceless' plosives. A significant main effect for *Place of Articulation* (labial vs. coronal; β=-16.58, *SE*=4.823, *t*=-3.437, *p*<.001) shows that the bilingual children produced /p/ with shorter VOT than /t/. No other main effects or interactions were significant.

Sub-model 'voiced' plosives: The only significant effect detected by the model was that of *Place of Articulation* (β=0.607, *SE*=0.222, *z*=2.739, *p*=.006), which shows that the bilingual children produced /b/ with prevoicing more frequently than /d/. No other main effects or interactions were significant.

Monolinguals

Main model. The main model detected a significant main effect for *Voicing* (β=29.966, *SE*=3.27, *t*=9.161, *p*<.001), which shows that the monolingual children produced longer VOT for 'voiceless' plosives than for 'voiced' plosives. A significant main effect for *Language* (β=25.203, *SE*=3.323, *t*=9.161, *p*<.001) confirms that monolingual German children produced longer VOT than monolingual Dutch children.

In addition, the main model detected a significant main effect for *Place of Articulation* (labial vs. coronal; β=-11.412, *SE*=2.773, *t*=-4.115, *p*<.001) and a three-way interaction between *Voicing* * *Language* * *Place of Articulation* (labial vs.

coronal; *β*=-7.422, *SE*=2.773, *t*=-2.676, *p*=.007). Post hoc analyses conducted on the data split by voicing and language investigated this interaction and found that the monolingual German children produced longer VOT in coronal than in labial plosives (voiced: *β*=-4.208, *SE*=1.851, *t*=-2.273, *p*=.023; voiceless: *β*=-23.545, *SE*=6.179, *t*=-3.810, *p*<.001), while no effect of *Place of Articulation* was detected in the VOT productions of the monolingual Dutch children (voiced: *β*=-16.545, *SE*=8.99, *t*=-1.84, *p*=.066; voiceless: *β*=-4.402, *SE*=3.242, *t*=-1.358, *p*=.174).

The main model furthermore detected significant interactions between *Voicing * Word Length* (*β*=-4.627, *SE*=1.432, *t*=-3.231, *p*=.001) and between *Language * Word Length* (*β*=-3.493, *SE*=1.432, *t*=-2.440, *p*=.015). Again, these interactions were further investigated by post hoc tests based on split data. Analyses on the data split by voicing suggest that both Dutch and German monolingual children produced longer VOT for 'voiceless' plosives when they occurred in monosyllabic words than in disyllabic words (*β*=-3.934, *SE*=1.371, *t*=-2.869, *p*=.004), but shorter VOT for 'voiced' plosives when they occurred in monosyllabic words than in disyllabic words (*β*=5.309, *SE*=2.591, *t*=2.049, *p*=.040). Analyses on the data split by language suggest that there is a non-significant trend towards longer VOT in disyllabic words in Dutch (*β*=-2.748, *SE*=1.641, *t*=-1.675, *p*=.094), and a non-significant trend towards longer VOT in monosyllabic words in German (*β*=4.176, *SE*=2.332, *t*=1.791, *p*=.073). No other main effects or interactions were significant.

Sub-model 'voiceless' plosives: A significant main effect for *Language* (*β*=31.679, *SE*=3.583, *t*=8.841, *p*<.001) shows that monolingual German children produce 'voiceless' plosives with longer VOT than monolingual Dutch children. The model detected a significant main effect for *Place of Articulation* (labial vs. coronal; *β*=-13.996, *SE*=3.502, *t*=-3.996, *p*<.001), and an interaction between *Language * Place of Articulation* (labial vs. coronal; *β*=-9.423, *SE*=3.502, *t*=-2.691, *p*=.007). This interaction was investigated in post hoc analyses conducted on the data split by language, and showed that only monolingual German children produced short VOT in /p/ than in /t/ (*β*=-23.545, *SE*=6.181, *t*=-3.810, p<.001), while this effect was non-significant for monolingual Dutch children (*β*=-4.402, *SE*=3.242, *t*=-1.358, *p*=.175). Moreover, the model detected a significant main effect for *Word Length* (*β*=-3.934, *SE*=1.372, *t*=-2.868, *p*=.004), as well as an interaction between *Language * Word Length* (*β*=-4.004, *SE*=1.372, *t*=-2.920, *p*=.004). This interaction was investigated in post hoc analyses conducted on the data split by language, and showed that only monolingual German children produced 'voiceless' plosives with longer VOT in monosyllabic words than in disyllabic words (*β*=-

8.092, *SE*=2.411, *t*=-3.356, *p*<.001), but this effect was non-significant for monolingual Dutch children (*β*=0.018, *SE*=1.275, *t*=0.014, *p*>.250). No other main effects or interactions were significant.

Sub-model 'voiced' plosives: The model detected a significant main effect for *Language* (*β*=-1.887, *SE*=0.285, *z*=-6.629, *p*<.001), which shows that Dutch monolingual children produced more 'voiced' plosives with prevoicing than German monolingual children. In addition, the model detected a significant main effect for *Place of Articulation* (*β*=1.030, *SE*=0.224, *z*=4.592, *p*<.001), showing that the Dutch and German monolingual children produced /b/ with prevoicing more frequently than /d/. No other main effects or interactions were significant.

Chapter 7

General discussion

The overarching goal of this dissertation was to draw a comprehensive picture of phonological acquisition of simultaneous bilingual preschoolers who are heritage speakers of German in the Netherlands. The focus had been set on preschool-aged heritage speakers for two primary reasons. First, preschoolers have a robust lexicon in which they presumably built phonological representations (Fikkert, 2010). Second, simultaneous bilingual preschoolers widely differ in the quantity of their heritage language input (e.g., Valdés, 2000a, 2000b). Individual differences in the quantity of heritage language input provide a suitable starting point for testing the influence of language input on bilingual children's phonological acquisition.

The comprehensive picture of bilingual preschoolers' phonological acquisition was drawn by empirical investigation of VOT production, perception, and input in Dutch-German simultaneous bilingual children aged 3;6 to 6;0 years. In parallel, I tested VOT production, perception and input in age-matched monolingual children acquiring either Dutch or German to identify possible divergences between bilingual and monolingual phonological acquisition. In this final chapter, the main findings of the research conducted for this dissertation are summarized and discussed.

Two general conclusions can be drawn from the research reported in **Chapter 2** to **Chapter 6**: 1) simultaneous bilingual children have language-specific, but not monolingual-like phonological systems, and 2) input quantity and quality direct the acquisition process of simultaneous bilingual children. These conclusions have important implications for the field of bilingual first language acquisition and provide a basis for a needed empirically-grounded theoretical model of simultaneous bilingual (phonological) acquisition (see section 4 of this chapter). Moreover, the research reported in this dissertation provides methodological insight into bilingualism research: different methods used to study simultaneous bilingual acquisition can reveal different perspectives into the nature of bilingual acquisition, and the phonetic correlates of the investigated contrast or phonemic segment are pivotal. In the following, I discuss and integrate the main results of chapters 2 through 6. I then discuss how the present results can shape future research and advance the development of a theoretical model targeting bilingual children's speech development.

1 Bilingual children's partially language-specific VOT production and language-specific VOT perception

The primary research questions addressed in chapters 2 through 4 were whether bilingual preschoolers' VOT production and VOT perception provide evidence for language-specific phonological systems, and whether these phonological systems are monolingual-like. The results of **Chapter 2** suggest that Dutch and German 'voiced' plosives interact in production, and the results of **Chapter 4** suggest that this interaction may occur at the level of features. The combined results of **Chapter 2**, **Chapter 3**, and **Chapter 4** provide evidence for language-specific phonological systems, which partially diverge from the phonological systems of monolingual children.

1.1 VOT production

In **Chapter 2**, the focus was on bilingual and monolingual children's production of VOT in the word-initial singleton plosives /p t k/ and /b d/ in Dutch and German. Bilingual children produced VOT differently from their monolingual peers in both languages. Differences between the VOT productions of bilingual and monolingual children had been observed in the past, but appeared to be limited to the heritage language (Deuchar & Clark, 1996; Johnson & Wilson, 2002; Kehoe, Lleó & Rakow, 2004; Khattab, 2000, 2003; Mayr & Siddika, 2016). Given that the heritage languages in previous research were prevoicing languages, it remained inconclusive whether VOT in the heritage language or rather VOT in the prevoicing language is prone to differential acquisition. Unlike previous research, the present research investigated VOT in a heritage language with aspiration and a majority language with prevoicing.

The production study reported in **Chapter 2** shows that in the majority language Dutch, bilingual children produced VOT for 'voiceless' plosives similar to their monolingual peers, but they produced fewer 'voiced' plosives with prevoicing in comparison to monolinguals. However, even monolingual Dutch children diverged from the adult-like production pattern and did not yet produce (almost) all 'voiced' plosives with prevoicing (Deighton-Van Witsen, 1976; Lisker & Abramson, 1964; van Alphen & Smits, 2004). The results of **Chapter 2** show that prevoicing of the bilingual children's majority language is also prone to differential acquisition. Given that prevoicing is typically acquired late even in monolingual acquisition (Khattab, 2000; Macken & Barton, 1980; MacLeod, 2016), it may pose a special challenge to bilingual children even in the majority language.

In the heritage language German, the bilingual children produced 'voiceless' plosives with shorter aspiration than monolingual children. The research presented in **Chapter 2** therefore provides evidence that aspiration is prone to differential acquisition in a heritage language acquisition context. Moreover, an unexpected finding was that the bilingual children produced prevoicing in German 'voiced' plosives. In German, 'voiced' plosives typically have short lag VOT, but prevoicing appears to be idiosyncratic in the speech of German-speaking adults (e.g., Jessen, 1998). Monolingual German-speaking children almost never produce prevoicing (Kehoe et al., 2004), most likely because it requires complex velopharyngeal adjustments that are typically acquired late by children speaking a prevoicing language (Khattab, 2000; Macken & Barton, 1980; MacLeod, 2016). Section 2 of this chapter elaborates on the bilingual children's production of prevoicing in German.

While **Chapter 2** focused on plosives that exist in both Dutch and German, **Chapter 4** was concerned with productions of the 'voiced' dorsal plosive /g/, which exists in German, but not in Dutch. When Dutch-German bilingual children were prompted to produce /g/ in names and a loan word in a Dutch context, they followed the same production pattern as monolingual Dutch-speaking children. I suggested that bilingual children made use of their Dutch phonological system when producing the non-native segment /g/. Moreover, despite the lack of /g/ in the Dutch phoneme inventory, bilingual children's productions of /g/ in German appeared to be influenced by their Dutch phonological system. This influence may occur directly from the Dutch voicing feature to German, which would suggest that cross-linguistic influence is not limited to interactions between specific segments. Alternatively, the influence from Dutch to German may occur indirectly through the children's German phonological system: presuming that the bilingual children's German /b/ and /d/ are affected by cross-linguistic influence from Dutch, feature generalization within their German phonological system can account for their prevoiced productions of [g] in German. In sum, the findings of **Chapter 4** suggest that features play an important role in children's speech.

1.2 VOT perception

Regarding the question whether bilingual children *produce* VOT language-specifically, the results of **Chapter 2** revealed ambiguous results: the bilingual children produced language-specific VOT for 'voiceless' plosives, but seemed to produce 'voiced' plosives with indistinguishable VOT in Dutch and German. In

Chapter 3, I followed up on the finding that bilingual children appear to have one merged category for Dutch and German 'voiced' plosives, which was observed in **Chapter 2** as well as in previous studies examining different language combinations (Deuchar & Clark, 1996; Johnson & Wilson, 2002 (the older child); Khattab, 2000; Mayr & Siddika, 2016). Measures of voicing perception revealed that bilingual children associate 'voiced' plosives with longer VOT in German than in Dutch. This result supports the hypothesis put forward in **Chapter 2** that the voicing systems of simultaneous bilingual children are language-specific, but that this cannot always be measured in their speech production.

The cumulative findings of **Chapter 2** and **Chapter 3** highlight the importance of combining speech production and speech perception methods to understand linguistic systems of bilingual children. Considering that much of the research on bilingual preschoolers' phonological acquisition is based on speech production measures (but see Brasileiro, 2009, for speech perception, and McCarthy, Mahon, Rosen & Evans, 2014, for a combination of speech production and speech perception in sequential bilingual preschoolers), bilingual children's phonological systems may in fact be more differentiated than previously assumed.

2 The phonetic correlates of the contrast matter

The experiments reported in **Chapter 2** and **Chapter 5** show an intriguing pattern: simultaneous bilingual children and late bilingual adults consistently seem to keep their 'voiceless' plosives apart in production in Dutch and German, but often produce 'voiced' plosives with the Dutch-like realization in both languages. Previous studies that took into account 'voiced' plosives showed a similar pattern (for children: Deuchar & Clark, 1996; Johnson & Wilson, 2002 (the older child); Khattab, 2000; Mayr & Siddika, 2016; for adults: Flege & Eefting, 1987a; Hazan & Boulakia, 1993; MacLeod & Stoel-Gammon, 2009; Mayr, Price & Mennen, 2012; Simon & Leuschner, 2010; Sundara, Polka & Baum, 2006). Moreover, the input quantity of the heritage language German was positively associated with target-like VOT production in German 'voiceless' plosives, but no association was observed for 'voiced' plosives in German. Similarly, as demonstrated in **Chapter 6**, maternal VOT was positively associated with children's VOT productions of Dutch and German 'voiceless' plosives, but no such association was detected for productions of 'voiced' plosives.

Based on these findings, I suggest that bilinguals' phonological acquisition follows an economic pattern that is based on the phonetic correlates of the voicing contrast. Acquiring target-like ‘voiceless’ plosives can be regarded as essential for both languages: first, producing ‘voiceless’ plosives without aspiration in German can lead to confusion, as the resulting sound signal will likely be perceived as the ‘voiced’ counterpart phoneme. Second, production of aspiration in Dutch may be perceived as foreign-accented (Flege, 1984; Flege & Eefting, 1987b; Major, 1987; Riney & Takagi, 1999; Sancier & Fowler, 1997; Schoonmaker-Gates, 2015). By contrast, productions of ‘voiced’ plosives with or without prevoicing in German is not lexically distinctive, which is also the case for English. Based on findings from English, it is likely that German listeners do not perceive productions of prevoicing as foreign-accented (Hazan & Boulakia, 1993).

In sum, there is a clear communicative benefit of keeping Dutch and German ‘voiceless’ plosives apart in production. This is presumably not the case for ‘voiced’ plosives, provided they are produced with prevoicing and do thus not interfere with Dutch ‘voiceless’ short lag plosives. For this reason, it appears to be more economic for bilinguals to merge their ‘voiced’ plosives in production.

Furthermore, the findings of **Chapter 2** and **Chapter 5** propose that instead of describing plosives in their phonological terms as ‘voiced’ or ‘voiceless’, their phonetic correlates appear to be crucial in describing and understanding bilingual children’s and bilingual adults’ VOT production. In **Chapter 2**, I found that bilingual children produced target short lag plosives, that is Dutch ‘voiceless’ and German ‘voiced’ plosives[19] in a monolingual-like manner, while they diverged from their monolingual peers in their production of target prevoiced and aspirated plosives. A similar pattern was observed in the parents’ VOT production as reported in **Chapter 5**: the native German parents acquired native-like short lag ‘voiceless’ plosives in Dutch (the L2) and maintained monolingual-like short lag ‘voiced’ plosives in German (the L1). On the contrary, they did not fully acquire Dutch prevoicing and did not fully maintain German aspiration.

These observed similarities in the acquisition and maintenance of short lag VOT in Dutch ‘voiceless’ plosives and German ‘voiced’ plosives have implications for the phonological representations of the voicing contrast in Dutch and German. The production results presented in this dissertation support the view that the phonetic correlates of the voicing contrast rather than the phonological status of plosives can account for the productions of VOT by the participants tested in this dissertation. These phonetic correlates of the voicing contrast can be represented through multiple monovalent features (e.g., Iverson & Salmons, 1995, 2003): Dutch

[19] This is in reference to the productions of devoiced ‘voiced’ plosives in German.

carries the feature [voice] for (prevoiced) ‘voiced’ plosives, while (short lag) ‘voiceless’ plosives remain unspecified [ ]. In German, (short lag) ‘voiced’ plosives are unspecified [ ], whereas (aspirated) ‘voiceless’ plosives carry the feature [spread glottis]. This multiple feature hypothesis can explain the observed similar production patterns for short lag plosives, be they phonologically ‘voiceless’ as in Dutch, or ‘voiced’ as in German.

3 Input quantity *and* input quality predict bilingual children’s speech production

The experiments reported in **Chapter 2, Chapter 5** and **Chapter 6** revealed important findings on the role of input quantity and quality in simultaneous bilingual acquisition. The importance of acknowledging language exposure patterns in bilingual acquisition has previously been highlighted (e.g., De Houwer, 2011; Gathercole, Thomas, Roberts, Hughes & Hughes, 2013). To this day, there appears to be little consistency in *how* input quantity is measured. It used to be common to ask parents for estimates of their children’s proportional input quantity, and this estimate was then used as a variable in analyses (e.g., Hoff, Core, Place, Rumiche, Señor & Parra, 2012; Pearson, Fernández, Lewedeg & Oller, 1997). By now, there are systematic and freely available tools that allow for precise assessments of language exposure, such as the Bilingual Language Experience Calculator (BiLEC; Unsworth, 2013). The BiLEC bases its input quantity calculations on information such as with whom the child is spending time on an average week day or weekend day, and which language this person speaks to the child. In addition, information on daycare attendance and extracurricular activities and the associated languages are inquired. In this dissertation, input quantity for each language has been calculated with the BiLEC and the resulting individual language exposure values (in percent) have been used as independent variable. In line with previous research that observed positive associations between the input quantity and bilingual children’s lexical knowledge, the results of this dissertation show that the input can furthermore directly influence fine phonetic detail in bilingual children’s speech.

In **Chapter 2**, I found that the quantity of the input to the heritage language German is positively associated with more target-like VOT production of German ‘voiceless’ plosives. Importantly, heritage language input quantity was not detectably associated with the bilingual children’s VOT production in the majority language Dutch. These findings suggest that bilingual children’s phonological acquisition in the heritage language benefits from more exposure to the heritage language, while it has no measureable negative consequences for the majority language.

These findings have important implications for parents and educators. It is still common that educators advise parents of bilingual children to avoid speaking the heritage language to their bilingual child (De Houwer, in press). Anecdotally, such advise had also been spontaneously reported by one of the bilingual families who participated in the research reported in this dissertation. Instead, the positive influence of heritage language exposure should be stressed. In addition to the known positive association between input quantity and bilingual children's lexical knowledge (Pearson et al., 1997), the study reported in **Chapter 2** is the first to show that the quantity of the input to the heritage language is associated with language-specific phonetic characteristics in the speech of bilingual children.

In addition, the input quality, namely non-native input, has often been assumed to play a role in bilingual children's language acquisition, but small sample sizes did not allow for statistically testing this hypothesis (Deuchar & Clark, 1996; Khattab, 2003; Klinger, 1962; Mayr & Montanari, 2015). In **Chapter 5**, I addressed VOT production of the bilingual children's parents and showed that the Dutch-native parents maintained monolingual-like VOT in Dutch, but strongly deviated from native-like VOT in German. The German-native parents produced non-native VOT in their L2 Dutch *and* in their L1 German. Specifically, the German mothers' VOT of 'voiceless' plosives in their L2 Dutch was longer (i.e., more German-like) in comparison to the VOT of monolingual Dutch mothers. Moreover, the German mothers produced 'voiceless' plosives in their L1 German with shorter (i.e., more Dutch-like) VOT than monolingual German mothers. A similar pattern was present in the bilingual children's VOT production, as reported in **Chapter 2.**

In **Chapter 6**, I connected the dots by showing that the differential VOT productions of the German-native mothers are associated with their children's VOT productions at an individual level in both Dutch and German. Recently, non-native VOT productions have been found in infant directed speech of bilingual parents addressed towards their bilingual infants, and the authors highlighted that these findings may explain the often-observed differences between bilingual and monolingual phonological development (Fish, García-Sierra, Ramírez-Esparza & Kuhl, 2017). The results of **Chapter 6** represent the first statistical evidence that fine phonetic detail in the speech input has measurable effects on bilingual children's speech production. Differences in the input of bilingual versus monolingual children thus contribute to differences in the speech between bilingual and monolingual children.

It is important to stress that the observed association between differential input and bilingual children's differential speech production does not mean that parents should avoid speaking their second language when their child is present. The finding rather highlights that the input shapes a child's language acquisition in many aspects and that differential acquisition outcomes are not exclusively attributed to cross-linguistic interaction within the (simultaneous) bilingual mind. Bilingual first language acquisition is a different scenario from monolingual first language acquisition, including different input. These different scenarios may be characterized by a different pace of acquisition, be it slower (e.g., slower acquisition of prevoicing in Dutch and aspiration in German) or faster (e.g., more productions of prevoicing in German) than monolingual acquisition. However, awareness must be raised to the fact that the observed differences between monolingual and bilingual children as well as between monolingual and bilingual adults were small phonetic differences. These differences may or may not have implications in real life. It remains a task for future research to determine whether native listeners detect such small phonetic differences, process them less efficiently, or associate them with foreign-accentedness.

4 Future directions: towards a model of simultaneous bilingual children's speech production

Since the seminal work of Ronjat (1913), the field of early bilingualism research has progressed tremendously. Yet, to this day, the subfield of phonetic-phonological aspects of the speech development in simultaneous bilingual children suffers from the lack of a dedicated empirically-grounded theoretical model (Hambly, Wren, McLeod & Roulstone, 2013; Kehoe, 2015). Due to this lack, partially compatible models on second language acquisition, such as the Speech Learning Model (SLM, Flege, 1995) are often used in the context of simultaneous bilingual acquisition (Fabiano-Smith & Bunta, 2012; Fabiano-Smith & Goldstein, 2010; Gildersleeve-Neumann & Wright, 2010). In **Chapter 2**, I furthermore consulted the A(rticulatory)-Map model on monolingual children's speech production to interpret bilingual children's production of VOT (McAllister Byun, Inkelas & Rose, 2016).

Both the SLM and A-Map have limitations when it comes to their application in a simultaneous bilingual acquisition context. The general assumption of the SLM that two languages co-occur in one shared phonological space, is likewise applicable to simultaneous bilingual acquisition. However, the SLM specifically assumes that the L1 phonological system is in place when the L2 is acquired. The L2 phonetic-phonological system is then acquired on top of the L1 system. According to the SLM, new L2 categories can be established when they are distinct from already existing L1 categories, and may otherwise be merged with already existing categories. These scenarios do not apply to a simultaneous acquisition context, in which two native languages are acquired in parallel from birth. Moreover, the SLM is designed to account for ultimate attainment in phonetic learning, and therefore does not take into account the speech development that bilingual children undergo. Importantly, Flege (2005) points out that the SLM should be used with participants whose language backgrounds allow to address the upper limits of L2 phonetic learning. As outlined above, simultaneous bilingual children cannot be considered as the target population the SLM seeks to describe. Moreover, Flege (2005) explicitly highlights that it is unreasonable to expect that L2 learners produce an L2 segment correctly when they are frequently exposed to inaccurate productions of this segment, suggesting that acquisition outcomes should always be related to the input. As established in **Chapter 5**, bilingual children are indeed exposed to inaccurate productions of VOT through their parents, which underlines the need for an empirically-based theoretical model specifically designed to account for simultaneous bilingual children's speech production.

The A-Map model explains differences between the speech of monolingual children and adults through constraints based on children's developing anatomy and motor-control. A child's speech production results from two competing forces, namely a child's pressure to match adult productions as well as the pressure to achieve reliable phonetic realizations, even if these do not match the adult-target form. The A-Map model further proposes that experience-based information on previous articulator movements and the resulting acoustic outputs are stored in episodic memory. With increasing experience, monolingual children become more precise in their speech production, until they finally match the adult-target.

In **Chapter 2** of this dissertation, I proposed that the experience component in the A-Map model offers a promising approach to account for the speech production in bilingual children. Based on a child's accumulated experience with a certain segment, the model can also explain slower or faster acquisition in bilingual children compared to monolinguals. However, the A-Map model would need to be extended in order to successfully account for bilingual children's speech production.

To be applicable to bilingual children's speech production, a model on bilingual children's phonological acquisition can be built on combined components of the SLM and the A-Map model. Such a model needs to take into account the SLM's assumption that two languages are accommodated in a shared phonetic space, and that these two languages may interact. Moreover, the A-Map model's production experience component as well as the constraints regarding children's anatomy and motor-control must be taken into account in a model designed for bilingual children's speech production. Regarding production experience, it is essential to include measures of bilingual children's language exposure into the model. Information on a bilingual child's language exposure combined with information on the monolingual acquisition trajectory for the segment under investigation allow for estimating the age at which bilingual children can be expected to master correct production of a segment or contrast. Information regarding proportional language exposure to the two languages may also allow for predictions regarding cross-linguistic interactions: if the proportional exposure to one language is very low, cross-linguistic interactions are more likely to occur. Moreover, as the adult-target in bilingual children's acquisition does not necessarily match the *monolingual*-adult-target, a model necessarily needs to take into account the speech input of bilingual children that is provided through their primary caretakers.

In sum, although this dissertation did not aim at designing a model on simultaneous bilingual children's phonological acquisition, it identified important characteristics that a future model should comprise. I propose that a future model should take into account the following questions:

1) What is the proportional distribution of the input per language?
2) What is the monolingual acquisition trajectory for the contrast/segment under investigation?
3) Is the child (primarily) exposed to non-native or attrited input?
4) Are there possible articulatory constraints that mask target-like production?

Question 4 is already provided by the A-Map model, and I suggest that when the answer to this question is "yes", speech perception measures should be administered before conclusions on the effect of bilingualism on a child's phonological development are drawn.

The answers to the four questions outlined above appear to be relevant based on the findings of this dissertation, but by no means are they exhaustive. It will be our future task to develop an empirically-grounded theoretical model of simultaneous bilingual first language acquisition, which will eventually facilitate generalizations across studies.

5 Concluding remarks

The simultaneous bilingual preschoolers examined in this dissertation grew up as heritage speakers of German in the Netherlands, and acquired German from one (and sometimes both) parents. All bilingual children communicated with ease in both Dutch and German, but they partially differed from monolingual children in their production of VOT. Speech production methods are commonly used when studying bilingual children, and researchers often observe that bilingual children differ from monolingual children in phonetic aspects of their speech. In this dissertation, I complemented speech production measures with a speech sound perception task. While bilingual children's production of VOT did not allow to draw firm conclusions about language-specific phonological systems, the speech sound perception task showed that bilingual children's voicing systems are language-specific, although this is not always measureable in their speech production. Moreover, the research presented in this dissertation empirically investigated the input of bilingual children. The results represent novel evidence that both the quantity of exposure to the heritage language and the specific phonetic realization of VOT in the speech of bilingual children's mothers are associated with bilingual children's VOT productions.

References

Brasileiro, I. (2009). *The effects of bilingualism on children's perception of speech sounds.* Utrecht: LOT.

De Houwer, A. (in press). Input, context and early child bilingualism: Implications for clinical practice. In A. Bar-On & D. Ravid (Eds.), *Handbook of Communication Disorders: Theoretical, Empirical, and Applied Linguistic Perspectives*. Berlin: Mouton de Gruyter.

De Houwer, A. (2011). Language input environments and language development in bilingual acquisition. *Applied Linguistics Review, 2*, 221–240.

Deighton-Van Witsen, H. V. (1976). Preliminary investigation concerning differences between voice onset time values of speakers of standard Dutch and of one of the dialects spoken in the province of Groningen (second version). In Institute of Phonetic Sciences (Ed.), *Proceedings of the Institute of Phonetic Sciences, 4* (pp. 113–133). Retrieved from http://www.fon.hum.uva.nl/archive/1976/Proc4Deighton.pdf

Deuchar, M., & Clark, A. (1996). Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics, 24,* 351–365.

Fabiano-Smith, L., & Bunta, F. (2012). Voice onset time of voiceless bilabial and velar stops in 3-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics & Phonetics, 26*, 148–163.

Fabiano-Smith, L., & Goldstein, B. (2010). Phonological acquisition in bilingual Spanish-English speaking children. *Journal of Speech, Language, and Hearing Research, 53*, 1–19.

Fikkert, P. (2010). Developing representations and the emergence of phonology: Evidence from perception and production. In C. Fougeron, B. Kuehnert, M. Imperio & N. Vallee (Eds.), *Laboratory Phonology 10* (pp. 227–260). Berlin: De Gruyter Mouton.

Fish, M. S., García-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. (2017). Infant-directed speech in English and Spanish: Assessments of monolingual and bilingual caregiver VOT. *Journal of Phonetics, 63*, 19–34.

Flege, J. E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America, 76,* 692–707.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Flege, J. E. (2005). The origins and development of the Speech Learning Model. Presented at the First Acoustical Society of America Workshop on L2 Speech Learning, Vancouver.

Flege, J. E., & Eefting, W. (1987a). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics, 15*, 67–83.

Flege, J. E., & Eefting, W. (1987b). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication, 6*, 185–202.

Gathercole, V., Thomas, E., Roberts, E., Hughes, C., & Hughes, E. (2013). Why assessment needs to take exposure into account: Vocabulary and grammatical abilities in bilingual children. In V. Gathercole (Ed.), *Issues in the assessment of bilinguals* (pp. 20–55). Bristol: Multilingual Matters.

Gildersleeve-Neumann, C., &Wright, K. (2010). English speech acquisition in 3- to 5-year-old children learning Russian and English. *Language, Speech, and Hearing Services in Schools, 41*, 429–444.

Hambly, H., Wren, Y., McLeod, S., & Roulstone, S. (2013). The influence of bilingualism on speech production: A systematic review. *International Journal of Language & Communication Disorders, 48*, 1–24.

Hazan, V. L., & Boulakia, G. (1993). Perception and production of a voicing contrast by French–English bilinguals. *Language and Speech, 36*, 17–38.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language, 39*, 1–27.

Iverson, G. K., & Salmons, J. C. (1995). Aspiration and laryngeal representation in Germanic. *Phonology, 12*, 369–396.

Iverson, G. K., & Salmons, J. C. (2003). Legacy specification in the laryngeal phonology of Dutch. *Journal of Germanic Linguistics, 15*, 1–26.

Jessen, M. (1998). *Phonetics and phonology of tense and lax obstruents in German*. Amsterdam: Benjamins.

Johnson, C., & Wilson, I. (2002). Phonetic evidence for early language differentiation: Research issues and some preliminary data. *International Journal of Bilingualism*, *6*, 271–289.

Kehoe, M. (2015). Cross-linguistic interaction: A retrospective and prospective view. In E. Babatsouli & D. Ingram (Eds.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2015* (pp. 141–167). Retrieved from http://ismbs.eu/data/documents/Proceedings-ISMBS-2015.pdf

Kehoe, M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition, 7*, 71–88.

Khattab, G. (2000). VOT in English and Arabic bilingual and monolingual children. In D. Nelson & P. Foulkes (Eds.), *Leeds Working Papers in Linguistics and Phonetics, 8* (pp. 95–122). Retrieved from http://www.leeds.ac.uk/arts/homepage/ 387/leeds_working_papers_in_linguistics_and_phonetics

Khattab, G. (2003). Age, input, and language mode factors in the acquisition of VOT by English-Arabic bilingual children. In M. J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the International Congress of Phonetic Sciences, 15* (pp. 3213–3216). Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_3213.pdf

Klinger, H. (1962). Imitated English cleft palate speech in a normal Spanish speaking child. *Journal of Speech and Hearing Disorders, 27*, 379–381.

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384–422.

Macken, M. A., & Barton, D. (1980). The acquisition of the voicing contrast in Spanish: A phonetic and phonological study of word-initial stop consonants. *Journal of Child Language, 7*, 433–458.

MacLeod, A. A. (2016). Phonetic and phonological perspectives on the acquisition of voice onset time by French-speaking children. *Clinical Linguistics & Phonetics, 30*, 584–598.

MacLeod, A. A., & Stoel-Gammon, C. (2009). The use of voice onset time by early bilinguals to distinguish homorganic stops in Canadian English and Canadian French. *Applied Psycholinguistics, 30*, 53–77.

Major, R. C. (1987). English voiceless stop production by speakers of Brazilian Portuguese. *Journal of Phonetics, 15*, 197–202.

Mayr, R., & Montanari, S. (2015). Cross-linguistic interaction in trilingual phonological development: The role of the input in the acquisition of the voicing contrast. *Journal of Child Language*, *42*, 1006–1035.

Mayr, R., Price, S., & Mennen, I. (2012). First language attrition in the speech of Dutch–English bilinguals: The case of monozygotic twin sisters. *Bilingualism: Language and Cognition, 15*, 687–700.

Mayr, R., & Siddika, A. (2016). Inter-generational transmission in a minority language setting: Stop consonant production by Bangladeshi heritage children and adults. *International Journal of Bilingualism*. Advance online publication. doi: 10.1177/1367006916672590.

McAllister Byun, T., Inkelas, S., & Rose, Y. (2016). The A-map model: Articulatory reliability in child-specific phonology. *Language, 92*, 141–178.

McCarthy, K. M., Mahon, M., Rosen, S., & Evans, B. G. (2014). Speech perception and production by sequential bilingual children: A longitudinal study of voice onset time acquisition. *Child Development, 85*, 1965–1980.

Pearson, B. Z., Fernández, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics, 18*, 41–58.

Riney, T., & Takagi, Y. (1999). Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning, 49*, 275–302.

Ronjat, J. (1913). *Le développement du langage observé chez un enfant bilingue*. Paris: Champion.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, 25*, 421–436.

Schoonmaker-Gates, E. (2015). On voice-onset time as a cue to foreign accent in Spanish: Native and nonnative perceptions. *Hispania, 98*, 779–791.

Simon, E., & Leuschner, T. (2010). Laryngeal systems in Dutch, English, and German: A contrastive phonological study on second and third language acquisition. *Journal of Germanic Linguistics, 22*, 403–424.

Sundara, M., Polka, L., & Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition, 9*, 97–114.

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*, 86–110.

Valdés, G. (2000a). Introduction. In Spanish for native speakers. *AATSP professional development series handbook for teachers K-16, Volume 1* (pp. 1–20). New York, NY: Harcourt College.

Valdés, G. (2000b). Teaching heritage languages: An introduction for Slavic language-teaching professionals. In O. Kagan & B. Rifkin (Eds.), *Learning and teaching of Slavic languages and cultures: Toward the 21st century* (pp. 375–403). Bloomington, IN: Slavica.

van Alphen, P. M., & Smits, R. (2004). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics, 32*, 455–491.

Samenvatting

Nina, Daan en Emma zijn alle drie vier jaar oud. Nina woont in Duitsland en haar moedertaal is Duits. Daan woont in Nederland en zijn moedertaal is Nederlands. Net als Daan groeit Emma op in Nederland, maar omdat Emma een Duitse moeder heeft, is zij tweetalig: ze spreekt zowel Duits als Nederlands. Op het eerste gezicht lijken Nina, Daan en Emma alleen van elkaar te verschillen in welke taal ze spreken én in de hoeveelheid talen die ze spreken. Maar, wat we niet moeten vergeten is dat Nina en Daan de hele dag alleen maar met moedertaalsprekers van respectievelijk het Duits en Nederlands omgaan terwijl Emma daarentegen zowel Nederlands als Duits van moedertaalsprekers (haar vader en moeder) hoort als Nederlands van een niet-moedertaalspreker. Haar moeder spreekt immers Nederlands als tweede taal. Emma moet haar aandacht dus verdelen over drie verschillende soorten taalinput. In dit proefschrift staan Nederlands-Duits tweetalige kinderen zoals Emma centraal. Alle Nederlands-Duits tweetalige kinderen die meededen aan de onderzoeken die we rapporteren in dit proefschrift hoorden op het moment dat ze meededen aan ons onderzoek gemiddeld meer Nederlands dan Duits, omdat alle kinderen op een Nederlandse kinderopvang of school zaten.

In dit proefschrift ligt de nadruk op de verwerving van het Nederlandse en Duitse klanksysteem van de tweetalige kinderen. De klanksystemen van beide talen hebben een aantal verschillende en een aantal overeenkomstige klanken. De Nederlandse diftong /œy/, zoals in <fr**ui**t> bestaat bijvoorbeeld niet in het Duitse klanksysteem en de Duitse /p͡f/, zoals in <**Pf**erd> (“paard”) bestaat juist niet in het Nederlandse klanksysteem. Tweetalige kinderen leren al vrij vroeg dat deze klanken slechts in een van de twee talen voorkomen.

Naast klanken die verschillen, zijn er ook een aantal klanken die in beide talen voorkomen. De manier waarop deze gemeenschappelijke klanken uitgesproken worden is wel subtiel verschillend. Bijvoorbeeld, zowel het Duits als het Nederlands kennen de /p t k/ en /b d/ klanken (ook wel “plosieven” genoemd), maar de manier waarop deze klanken uitgesproken worden, verschilt per taal. Een goed voorbeeld om dit verschil te illustreren is het Duitse én Nederlandse woordje <papa>/<Papa>. Om de Nederlandse [p] te produceren, moet je het volgende doen: je lippen sluiten, lucht in je mond laten ophopen en vervolgens je lippen openen waardoor de opgehoopte lucht ontsnapt en er een kleine explosie ontstaat. Gelijktijdig met het openen van je lippen, beginnen ook je stembanden te trillen zodat je de [a] kan produceren. Dit laatste, het gelijktijdig openen van de lippen en trillen van de stembanden, is een belangrijk detail en hierin verschilt de Nederlandse uitspraak van de Duitse uitspraak, zoals je hieronder kan lezen.

Wanneer Duitse moedertaalsprekers <Papa> uitspreken, trillen hun stembanden niet gelijktijdig met het openen van de lippen, maar pas iets later. Hierdoor komt er een klein pufje lucht vrij, wat aspiratie genoemd wordt. Het gevolg is dat het Duitse <Papa> voor Nederlanders klinkt als <Phapa>. Interessant is dat de Nederlandse uitspraak van <papa> voor Duitsers klinkt als <Baba> doordat Duitsers hun [b] uitspreken als een Nederlandse [p] (dus met het gelijktijdig trillen van de stembanden en openen van de lippen). Wanneer Nederlanders op hun beurt een [b] produceren, trillen de stembanden al voordat ze hun lippen van elkaar afhalen. Je kan het zelf voelen wanneer je een woord uitspreekt dat begint met een [b], bijvoorbeeld <beer>. Probeer maar eens. Wanneer je je hand op je keel legt, zal je de stembanden voelen trillen nog voordat je daadwerkelijk de [b] uitspreekt en je je lippen van elkaar afhaalt. Bovenstaande voorbeelden maken duidelijk dat Nederlands-Duits tweetalige kinderen 1) moeten ontdekken welke klanken alleen in het Nederlands voorkomen en welke klanken alleen in het Duits voorkomen en 2) moeten leren welke klanken in beide talen voorkomen, maar subtiel anders uitgesproken worden. In dit proefschrift heb ik samen met mijn collega’s onderzocht of tweetalige kinderen tussen de 3 en 6 jaar oud zich bewust zijn van deze verschillen tussen hun beide moedertalen.

In het eerste deel van het proefschrift (hoofdstuk 2) staat de uitspraak van Nederlands-Duits tweetalige kinderen centraal. Door middel van akoestische spraakanalyses ontdekten we dat de kinderen de [p t k] klanken anders uitspraken in het Nederlands dan in het Duits, wat betekent dat de kinderen zich bewust zijn van de fonetische verschillen tussen deze klanken in beide talen. Uit de akoestische spraakanalyse kwam naar voren dat de kinderen de [p t k] hetzelfde uitspraken als hun eentalige Nederlandse leeftijdsgenootjes. Dit valt waarschijnlijk te verklaren doordat ze naar Nederlandstalige kinderopvang of scholen gingen. In het Duits was de invloed van het Nederlands hoorbaar in hun uitspraak van de [p t k] klanken (“vernederlandst”): de tijdsduur tussen het openen van de lippen en het trillen van de stembanden (de aspiratie) was korter dan bij leeftijdsgenoten die eentalig Duits zijn opgevoed.

Daarnaast ontdekten we ook dat de tweetalige kinderen die relatief veel Duits hoorden, bijvoorbeeld omdat hun ouders meer Duits tegen ze spraken of omdat ze hun Duitse opa en oma regelmatig zagen, de [p t k] klanken Duitser uitspraken dan de tweetalige kinderen die relatief minder Duits hoorden. Maar misschien nog wel belangrijker, het was niet zo dat de Nederlandse uitspraak van deze kinderen die relatief meer Duits hoorden “verduitst” was. Het feit dat ze thuis meer Duits hoorden had dus geen (negatieve) invloed op de Nederlandse uitspraak, maar alleen een positieve invloed op hun Duitse uitspraak en daarom is een van onze conclusies dat meer blootstelling aan de thuistaal een positieve invloed heeft op de uitspraak van tweetalige Nederlands-Duitse kinderen.

Een verrassende vondst is dat dit verschil in uitspraak niet gevonden werd voor de [b d] klanken. De Nederlands-Duits tweetalige kinderen spraken deze twee klanken dus hetzelfde uit in het Nederlands als in het Duits. Hun uitspraak verschilde wel van die van eentalige Nederlandse én eentalige Duitse kinderen. Eigenlijk gebruikten ze een soort tussenvorm. Dit kan twee dingen betekenen: de tweetalige kinderen zijn zich (nog) niet bewust van het verschil tussen de Nederlandse en Duitse [b d] klanken of ze zijn zich er wel van bewust, maar gebruiken het zelf (nog) niet in hun uitspraak. De resultaten die we hier beschreven hebben, zijn ook opgepikt door Marc van Oostendorp en samengevat op zijn blog *Neerlandistiek* (van Oostendorp, 2017).

In hoofdstuk 3 gaan we dieper in op de vraag of de tweetalige kinderen het verschil tussen de Duitse [b d] en Nederlandse [b d] inderdaad niet horen en of dat kan verklaren waarom hun uitspraak van deze klanken in beide talen hetzelfde is. Om antwoord op deze vraag te krijgen, maakten we gebruik van zelf gesynthetiseerde /pa/ en /ba/ klanken. Er waren drie soorten klanken: 1) typisch Nederlandse /pa/ en /ba/ klanken; 2) typisch Duitse /pa/ en /ba/ klanken en 3) een tussenvariant van beide klanken. Wanneer we volwassenen naar deze tussenvarianten lieten luisteren, namen de moedertaalsprekers van het Nederlands de tussenvariant waar als /pa/ terwijl moedertaalsprekers van het Duits de tussenvariant waarnamen als /ba/. Om te achterhalen hoe de tweetalige kinderen deze tussenvariant waarnemen, ontwikkelden we een eigen computerspel. De kinderen speelden het spel twee keer: één keer in het Nederlands en één keer in het Duits. De resultaten lieten zien dat de kinderen de tussenvariant als /pa/ waarnamen wanneer ze het spel in het Nederlands speelden en als /ba/ wanneer ze het spel in het Duits speelden. Dit betekent dus dat tweetalige kinderen zich wel degelijk bewust zijn van het akoestische verschil tussen de Nederlandse /b/ en de Duitse /b/, ondanks dat ze het verschil in uitspraak op dit moment in hun taalontwikkeling (nog) niet realiseren.

In hoofdstuk 2 en 3 hebben we vastgesteld hoe Nederlands-Duits tweetalige kinderen plosieven in beide talen uitspreken en waarnemen. In hoofdstuk 4 gaan we nog een stapje verder en vragen we ons af wat voor soort (impliciete) kennis de kinderen over de eigenschappen van de klanken hebben. Naast de eerdergenoemde verschillen in uitspraak, is er nog een ander verschil dat het mogelijk maakt deze wat abstractere onderzoeksvraag te beantwoorden. Het Duits kent namelijk naast de drie stemloze plosieven /p t k/ ook drie stemhebbende plosieven /b d g/. Deze /g/ als in het Duitse woord <Garten> (‘tuin’) bestaat niet in het Nederlands en heeft in de manier van uitspreken twee eigenschappen gemeen met de /b/ en /d/: 1) de klank wordt geproduceerd bij het openen van de lippen en na een volledige sluiting van de lippen, hoewel in het geval van de /g/ de sprekers het achterste van hun tong tegen het zachte gehemelte drukken terwijl de lippen nog gesloten zijn, en 2) de klank is stemhebbend wat betekent dat de stembanden beginnen te trillen vlak voor of tijdens het openen van de lippen. Alle tweetalige kinderen in ons onderzoek kennen de /g/ klank en door middel van een spelletje onderzochten we hoe de kinderen deze /g/ uitspraken in een Nederlandse context. De kinderen werden voorgesteld aan drie fictieve karakters: een meisje met de naam <Gabi>, een jongen met de naam <Gero> en een hond die luisterde naar de naam <Gizmo>. De resultaten lieten zien dat de kinderen een andere /g/ klank gebruikten in de Nederlandse dan in de Duitse context. Dit is bijzonder, zeker omdat de kinderen in principe gewoon de Duitse /g/ hadden kunnen gebruiken, omdat deze klank in het Nederlands niet bestaat. Het feit dat de tweetalige kinderen dit niet deden en de klank toch anders uitspraken, laat zien dat ze hun kennis van het Nederlands gebruikten wat erop wijst dat tweetalige kinderen de twee talen ook als twee aparte talen hebben opgeslagen in hun brein.

In de laatste twee hoofdstukken van dit proefschrift (hoofdstuk 5 en hoofdstuk 6) werd de uitspraak van de ouders van de tweetalige kinderen onder de loep genomen. De meeste kinderen uit het onderzoek hadden één Nederlandstalige ouder en één Duitstalige ouder. De Nederlandstalige ouder sprak Duits als tweede taal en de Duitstalige ouder sprak Nederlands als tweede taal. In hoofdstuk 5 beschrijven we onze bevinding dat Duitstalige ouders, niet geheel onverwacht, de Nederlandse plosieven niet volledig op de Nederlandse manier uitspraken. Wel verrassend is de bevinding dat ze ook de Duitse plosieven niet meer volledig op de Duitse manier uitspraken. Dit komt waarschijnlijk doordat deze ouders al lange tijd in Nederland wonen en dus relatief weinig met moedertaalsprekers van het Duits omgaan. Met andere woorden, de uitspraak van Duitse plosieven bij Duitse ouders zijn vernederlandst en de uitspraak van de Nederlandse plosieven zijn verduitst. Bij de Nederlandse ouders met Duits als tweede taal zagen we dat alleen dat hun uitspraak van de Duitse plosieven wat vernederlandst was, hun uitspraak van de Nederlandse plosieven was gelijk aan die van Nederlandse ouders die geen Duits als tweede taal spreken.

Eerder onderzoek heeft aangetoond dat bij de taalontwikkeling van kinderen vooral de uitspraak van de moeder veel invloed heeft. In hoofdstuk 6 onderzoeken we daarom of de tweetalige kinderen de niet-moedertaal uitspraak van hun moeders (dus bijvoorbeeld de Duitse uitspraak van een Nederlandse moeder) overnemen. Om een antwoord te krijgen op deze vraag vergeleken we de uitspraak van de moeder met de uitspraak van haar kind in zowel het Nederlands als in het Duits. We toonden aan dat de uitspraak van de /p t k/ klanken van de kinderen hetzelfde was als de uitspraak van hun moeders, in zowel het Nederlands als in het Duits. Oftewel, wanneer de moeder deze klanken op een manier als typisch voor moedertaalsprekers van het Nederlands dan wel Duits uitsprak, dan deed haar kind dat ook. Het betekent ook dat wanneer een moeder de klanken op een atypische manier uitsprak, die atypische uitspraak ook werd overgenomen door haar kind. We zijn de eerste onderzoekers die laten zien dat een verschil in uitspraak tussen eentalige en tweetalige kinderen dus samenhangt met de input die zij van hun moeder krijgen.

Samenvattend, dit proefschrift laat zien dat de spraakontwikkeling van tweetalige kinderen meer gelijk is aan die van eentalige kinderen dan gedacht: door taalproductie- en taalperceptiematen te combineren hebben we laten zien dat het tweetalige kinderbrein taalspecifieke informatie over de klanken in beide talen bevat. Kleine verschillen in de uitspraak van tweetalige en eentalige kinderen hangen samen met 1) de totale tijd dat het kind een taal hoort en 2) de specifieke input die het kind krijgt: tweetalige kinderen hebben vaak ouders die hun tweede taal pas op latere leeftijd hebben geleerd. Het gevolg is dat de uitspraak van deze ouders anders is dan de uitspraak van een moedertaalspreker van de desbetreffende taal. Hieruit volgt automatisch dat het klankaanbod van tweetalige kinderen verschilt van het klankaanbod van eentalige kinderen. Om deze redenen moeten we niet te snel concluderen dat tweetalige kinderen, zoals Emma, verward zijn of achterliggen op eentalige kinderen zoals Daan en Nina wanneer hun uitspraak verschilt ten opzichte van deze eentalige kinderen. Nee, de uitspraak varieert omdat deze samenhangt met de uitspraak van hun moeder.

Referentie

van Oostendorp, M. (2017, juli 13). Nederlandse papa, Duitse phapha [Blog post]. Beschikbaar via http://www.neerlandistiek.nl/2017/07/nederlandse-papa-duitse-phapha/

Curriculum Vitae

Antje Stoehr was born on March 25th, 1987 in Bad Oldesloe in Northern Germany. She obtained a bachelor's degree in French Linguistics and Business Administration from the University of Hamburg (Germany). As an undergraduate student, she gained first research experience at the Research Center on Multilingualism. Funded by the German Academic Exchange Service (DAAD), Antje continued her studies at the University of Delaware (USA), where she graduated with a Master of Arts degree in Linguistics and Cognitive Science. After graduation, she commenced her PhD at Radboud University Nijmegen and the International Max Planck Research School for Language Sciences (the Netherlands). With a grant awarded by the Fulbright Foundation, she spent the third year of the PhD program at the Bilingualism and Language Development Laboratory at the Pennsylvania State University (USA). Antje will soon start working as a postdoctoral researcher at the Basque Center on Cognition, Brain and Language in Donostia-San Sebastián (Spain).

Publications

Thesis-related publications

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017). Second language attainment and first language attrition: The case of VOT in immersed Dutch-German late bilinguals. *Second Language Research, 33*, 483–518.

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017). Heritage language exposure impacts voice onset time of Dutch-German simultaneous bilingual preschoolers. *Bilingualism: Language and Cognition*. Advance online publication. DOI: 10.1017/S1366728917000116

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (under revision). Bilingual preschoolers' speech is associated with non-native maternal language input.

Other publications

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 the weaker and stronger language. *Journal of French Language Studies, 24*, 347–376.

Kupisch, T., Akpınar, D., & **Stöhr, A.** (2013). Gender assignment and gender agreement in adult bilinguals and second language learners of French. *Linguistic Approaches to Bilingualism, 3*, 150–179.

Stöhr, A., Akpınar, D., Bianchi, G., & Kupisch, T. (2012). Gender marking in Italian-German heritage speakers and L2-learners of German. In K. Braunmueller & C. Gabriel (Eds.), *Multilingual Individuals Multilingual Societies [Hamburg Studies on Multilingualism 13*] (pp. 153–170). Amsterdam: Benjamins.

Acknowledgements

We have seen in this dissertation that the input plays an important role in bilingual children's language development. Like the children in this dissertation, I received invaluable input during the past years, which had an incredibly important influence on my professional and personal development.

The academic input

Titia, Paula, and Janet: You have been the best supervisors I could have asked for. The quality and quantity of your input guided me through this PhD with the liberty to develop my own ideas, while providing me with the guidance I needed to make the right decisions along my way. Titia, I was your first PhD student, and your supervision excelled in every aspect. You taught me almost everything I know about phonetics and statistics, and you are also a master of life lessons and efficiency. It is because of you that I learned that everything is possible if only I try hard enough. If there is one supervisor on this planet who can increase the supervision quality after moving 15,000 km away, then I know for a fact that this is you. Your origami of the Australian wildlife will always have a special place on my desk! Paula, you always ensured that I would not lose track of the big picture. Whenever I felt like in a dead end and did not know where to go next, you were the one who reminded me that I can find the answer myself as soon as I know what I want to achieve. It is probably a once-in-a-lifetime experience to have a supervisor who invites the entire group to one-week writing retreats (which turned out to be so much more than that). Janet, the year I spent at Penn State was clearly one of the most exciting ones of this PhD! Your positive way of thinking and your clear words changed my personal view and made every aspect of this PhD a whole lot easier. It is you who taught me that I should never waste energy worrying about things that may or may not happen. Thank you so much for bringing new academic and personal perspectives into my research and life, and for being such a supportive and welcoming supervisor! Titia, Paula, and Janet, you have provided so much more than purely academic input to me over the past years and I truly cannot thank you enough!

My deepest gratitude goes to the members of the manuscript committee and doctoral examination board: Rob Schoonen, Mirjam Broersma, Annick De Houwer, Silke Hamann, Margaret Kehoe, Wander Lowie, and Ellen Simon. I profoundly appreciate that you allocated time in your busy schedules to evaluate this dissertation and to join my public defense.

The decision to get a PhD resulted from the great research experience I gained as an undergraduate student at the University of Hamburg. Tanja, you introduced me to the fascinating field of bilingualism research, and you have been my mentor ever since. Together with Joelle, Deniz, and all the other wonderful E11 members at the Research Center on Multilingualism, you showed me how much fun it is to do research. I truly cannot imagine where I would be today had I not met you approximately ten years ago. I am so glad we are still in touch, and I would be thrilled if our scientific paths crossed again in the future. Tausend Dank!

The research input

To my dearest fellow FLAdies and Maarten: It has been such a pleasure to join a group of colleagues who soon became friends! Christina, Evelien, Helen, Imme, Ingeborg, Laura, Maarten, Nadine, Nienke, Paula, Sho, Steffi, Susanne, Tineke, and Titia, you were the best colleagues, writing buddies, travel mates, and co-conference organizers I can ever imagine! Your input clearly shaped my research, and perhaps even more crucially, me as a researcher. I am so lucky to have had the opportunity to work aside you and to have individual memories with each of you.

During my research stay at Penn State University in the midst of Central Pennsylvania, I was surrounded by amazing colleagues in the Bilingualism and Language Development Lab and the Language Acquisition Lab. Thank you, Angela, Caitlin, Carla, Cindy, Chris, Dora, Fatemeh, Janet, Kaitlyn, Karen, Katherine, Megan, Pablo, Patricia, and Sarah. I would especially like to thank Fatemeh and Katherine for their help and great company in the EEG van, and Kaitlyn for her substantial help in getting my first EEG study off the ground. Jessica and Nelson, I cannot thank you enough for helping me test numerous participants. I would also like to thank my Penn State office mates Annie, Christian, Emilia, Grant, Janina, Kinsey, and Randi for making the Moore basement a brighter place! Janet, Frans, Merel, Rosa (and Bella!), thank you so much for your hospitality! The numerous get-togethers at your house and the delicious food enriched my stay in State College tremendously.

Throughout my PhD, many colleagues outside the First Language Acquisition group provided invaluable input to my research. Special thanks go to the colleagues of the Baby Research Centre Nijmegen (especially to Angela for helping me recruit many participants), the members of the Bilingualism and Cognition group (especially to Sharon for getting the group together and to Chantal, Kathleen, and Zeynep for organizing the meetings). Thank you to all CLS colleagues (especially to Margret for organizing so many lab lunches)! Thank you to the Utrecht colleagues in René Kager's group: Anna Sara, Brigitta and Liquan. Discussing my work with colleagues with such diverse backgrounds broadened my mind in so many aspects. Last but not least, thank you to all Nijmegen linguists for making Radboud University the best linguistics research institution in the world according to the Center for World University Rankings in the year 2017! It is a great honor to know for a fact that there could not have been a better place to get a PhD!

The practical and supportive input

This thesis would not have been possible without the practical and supportive input I received over the past years.

First and foremost, I would like to express my gratitude to the wonderful interns and student assistants who invested their brain power, time, and physical energy (sorry for the heavy bag!) to help me collect and annotate the data for this dissertation. Ann-Katrin, Arjan, Eva, Franziska, Jana, Lottie, Maaike, Natascha, and Rianne: Thank you so much! Without your help, I would certainly still be collecting data instead of writing these lines.

Laura, Imme, and Ingeborg, thanks a million for your help with the dissertation summary: Laura, thank you for your input on the first draft! Imme, thank you for your input and for translating the summary into Dutch! Ingeborg, thank you for the final fine-tuning!

Wencui and Mirjam, I am so thankful that you made your lovely dinosaur experiment available to me. I cannot even imagine how much time and effort you invested into the design of this beautiful experiment, and it was more than generous of you to let me use it for my research.

I would like to thank the colleagues of the University of Cologne and especially Frank Kirchhoff for letting me use their office space for testing German native speakers back in 2013.

The International Max Planck Research School for Language Sciences at the MPI for Psycholinguistics and the Graduate School for the Humanities at Radboud University have provided so much supportive input by ensuring that I made progress and was coping well with PhD student life. Dirkje, Els, Kevin, Peter, and Tanja, thank you so much for always being available and supportive. It meant a lot!

I would also like to express my gratitude to the MPI for letting me use their testing equipment and technical support. Ronald, thank you for all the time and effort you invested in adapting the dinosaur experiment for my needs.

Karin and Meggie, thank you for your kind support in publishing this book. You made the entire process a whole lot easier for me!

Numerous families spared hours of their time to participate in the experiments conducted for this dissertation. Without your curiosity, openness, and patience this dissertation could not have been written. Dank jullie wel & vielen lieben Dank!

The social input (and beyond)

Every aspect of PhD student life became so much more enjoyable thanks to the FLAdies and Maarten, Chara, Conny, Guusje, Hongling, Jule, Karin, Kathleen, Mijntje, Polina, Remy, Thordis, Wencui, and all other fellow IMPRSers. Without you, this entire PhD experience would not have been the same!

Christina, this is also the place to direct a very special Dankeschön to you. When I joined the FLA group, you became a friend immediately. Once we were introduced to each other, you made sure I found my way around both socially and academically, and you continue doing the same ever since. Thank you for being such a caring friend and inspirational colleague!

Chara and Imme, thank you so much for literally and figuratively standing by my side as my paranymphs. Your friendship and supportive spirits will certainly help me stay calm(er) during this exciting day!

I am more than grateful for the support I received from my family and friends from Ahrensburg, Hamburg, and Delaware – thank you so much for helping me move, visiting me, and sending me boxes full of Franzbrötchen. Eure Unterstützung bei allen bisherigen und zukünftigen Vorhaben bedeutet mir sehr viel!

Gunnar, there is no way of putting into words how much I appreciate all your input and support over the past years (ranging from cooking me dinner to reading every single page of this book). Because no words could do justice to what this means to me, see last page of this book.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children’s language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*

18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*

37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B. M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*

56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Conny de Vos*

73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*

74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*

75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*

76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*

77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*

78. Contributions of executive control to individual differences in word production. *Zeshu Shao*

79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*

80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*

81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*

82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*

83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*

84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*

85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*

86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*

87. The acquisition of morphophonological alternations across languages. *Helen Buckler*

88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*

89. Rediscovering a forgotten language. *Jiyoun Choi*

90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*
91. Infants'understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*
92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sümer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. lncrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*
101. Assessing birth language memory in young adoptees. *Wencui Zhou*
102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*
103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*
104. Conversation Electrified: The Electrophysiology of Spoken Speech Act Recognition. *Rósa Signý Gisladottir*
105. Modelling Multimodal Language Processing. *Alastair Smith*
106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*
107. Situational variation in non-native communication. *Huib Kouwenhoven*
108. Sustained attention in language production. *Suzanne Jongman*

109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntatic, and orthographic effects. *Malte Viebahn*
110. Nativeness, dominance, and the flexibility of listening to spoken language. *Laurence Bruggeman*
111. Semantic specificity of perception verbs in Maniq. *Ewelina Wnuk*
112. On the identification of FOXP2 gene enhancers and their role in brain development. *Martin Becker*
113. Events in language and thought: The case of serial verb constructions in Avatime. *Rebecca Defina*
114. Deciphering common and rare genetic effects on reading ability. *Amaia Carrión Castillo*
115. Music and language comprehension in the brain. *Richard Kunert*
116. Comprehending Comprehension: Insights from neuronal oscillations on the neuronal basis of language. *Nietzsche H. L. Lam*
117. The biology of variation in anatomical brain asymmetries. *Tulio Guadalupe*
118. Language processing in a conversation context. *Lotte Schoot*
119. Achieving mutual understanding in Argentine Sign Language. *Elizabeth Manrique*
120. Talking Sense: the behavioural and neural correlates of sound symbolism. *Gwilym Lockwood*
121. Getting under your skin: The role of perspective and simulation of experience in narrative comprehension. *Franziska Hartung*
122. Sensorimotor experience in speech perception. *Will Schuerman*
123. Explorations of beta-band neural oscillations during language comprehension: Sentence processing and beyond. *Ashley Lewis*
124. Influences on the magnitude of syntactic priming. *Evelien Heyselaar*
125. Lapse organization in interaction. *Elliott Hoey*
126. The processing of reduced word pronunciation variants by natives and foreign language learners: Evidence from French casual speech. *Sophie Brand*

127. The neighbors will tell you what to expect: Effects of aging and predictability on language processing. *Cornelia Moers*
128. The role of voice and word order in incremental sentence processing. *Sebastian Sauppe*
129. Learning from the (un)expected: Age and individual differences in statistical learning and perceptual learning in speech. *Thordis Neger*
130. Mental representations of Dutch regular morphologically complex neologisms. *Laura de Vaan*
131. Speech production, perception, and input of simultaneous bilingual preschoolers: Evidence from voice onset time. *Antje Stoehr*