

# Heritage language exposure impacts voice onset time of Dutch–German simultaneous bilingual preschoolers\*

ANTJE STOEHR

Centre for Language Studies, Radboud University, Nijmegen, The Netherlands and

International Max Planck Research School for Language Sciences, Nijmegen, The Netherlands

TITIA BENDERS

ARC Centre of Excellence in Cognition and its Disorders, North Ryde, Australia and

Department of Linguistics, Macquarie University, North Ryde, Australia

JANET G. VAN HELL

Department of Psychology and Center for Language Science, The Pennsylvania State University, University Park, PA, USA

PAULA FIKKERT

Centre for Language Studies, Radboud University, Nijmegen, The Netherlands

(Received: May 25, 2016; final revision received: November 11, 2016; accepted: January 18, 2017; first published online 22 June 2017)

*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.*

Keywords: child bilingualism, language exposure, heritage language, voice onset time (VOT), cross-linguistic influence

## 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 on mainly 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, published online 17 October, 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,

\* We thank Arjan Cuppen, Lottie Gort, Eva Koch, Jana Loh and Rianne Vlaar for their help collecting the data and Franziska Klier, Ann-Katrin Ohlerth and Natascha Roos for their help annotating the data. We are also grateful to three anonymous reviewers for insightful comments and suggestions on an earlier version of this paper.

Supplementary material can be found online at <http://dx.doi.org/10.1017/S1366728917000116>

Address for correspondence:

Antje Stoehr, Centre for Language Studies, Radboud University, PO Box 9103, 6500 HD Nijmegen, The Netherlands.  
[a.stoehr@let.ru.nl](mailto:a.stoehr@let.ru.nl)

the two languages are accommodated in their brain and are likely to influence each other, a phenomenon known as CROSS-LINGUISTIC INFLUENCE (CLI; Fabiano-Smith & Bunta, 2012; Fabiano & Goldstein, 2005; 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; Læufer, 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, published online June 22, 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 & Tsimpili, 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 native-like sounds in both of their languages.

The second hypothesis, henceforth the 'Equivalence Classification Hypothesis' (cf. 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 outputs 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, 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).

### *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 cord 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<sup>1</sup> 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 German and English implement the voicing contrast with short lag VOT vs. aspiration<sup>2</sup>. Language-specific VOT values within these ranges may differ cross-linguistically.

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 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,

<sup>1</sup> 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).

<sup>2</sup> Swedish distinguishes prevoiced 'voiced' and aspirated 'voiceless' plosives (Beckman, Helgason, McMurray & Ringen, 2011).

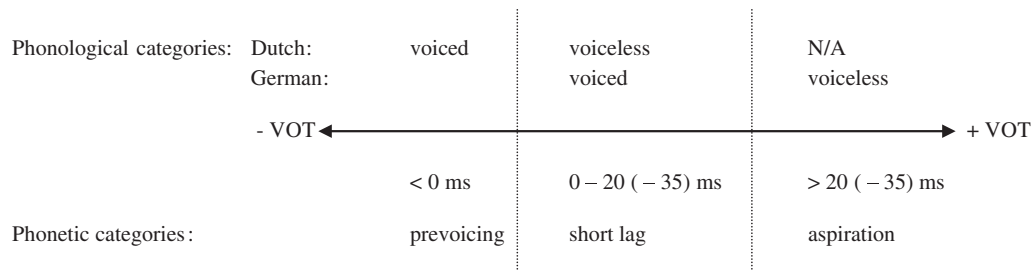


Figure 1. The VOT continuum: phonological and phonetic categories.

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).

#### 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, 2016).

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 prevoiced 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 prevoiced 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, 2016). Arabic-speaking

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

Voicing contrast	Study	Language	Age	<i>N</i>
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 & Koenigsnecht (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 (2016)	French	2;6–4;6	63
	Khattab (2000)	Arabic	5;4, 7;4, 10;3	3

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.

### *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, published online 17 October, 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. 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 (published online 17 October, 2016) who investigated VOT of twenty 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

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

Study	Prevoicing lang.	Aspiration lang.	Majority lang.	Age	<i>N</i>
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 (published online, 17 October, 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	English & Spanish	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.

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 (published online 17 October, 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, published online 17 October, 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.

Bilingual children can follow different patterns of VOT development even if their acquisition context is similar. 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 (cf. Johnson & Wilson, 2002; Mayr & Siddika, published online 17 October, 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 (cf. 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 above review 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 (cf. Mayr & Siddika, published online 17 October, 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 raises the question of 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.

### **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 yet have acquired prevoicing in Dutch and aspiration in German similarly to 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;6 and 6;0 years, older bilingual and monolingual children are expected to produce prevoicing and aspiration more reliably than younger children.

### **Method**

#### **Participants**

Eighty-eight children between 3;6 and 6;0 years participated in this study: 29 Dutch–German bilinguals ( $M_{\text{age}} = 4;7$ , range 3;7–5;11; 14 female), 30 Dutch monolinguals ( $M_{\text{age}} = 4;9$ , range 3;6–6;0; 17 female) and 29 German monolinguals ( $M_{\text{age}} = 4;8$ , range 3;6–6;0; 20 female)<sup>3</sup>. 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

<sup>3</sup> 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.

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 other child's first regular exposure to Dutch started at 0;6. 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,  $t(28) = 2.89$ ,  $p = .007$ . 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<sup>4</sup> (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.

<sup>4</sup> 0 = Kindergarten, 1 = Elementary school, 2 = 10<sup>th</sup> grade, 3 = High school, 4 = Bachelor's degree, 5 = Master's degree, 6 = Doctorate.

### Materials and procedure

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 were selected from the Dutch version of the MacArthur-Bates 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). Tables S1 and S2 in the online supplementary materials (Supplementary Material) provide an overview of the Dutch and German target words, respectively. 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 Table S3 in the online supplementary materials with descriptive statistics showing how the children's VOT differs by vocalic context. Table S3 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.

Testing took place in a quiet room at the children's homes. At the beginning of the session, parents gave informed consent and 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 children named all target words in two different picture-naming tasks to enhance the number of produced tokens per child while keeping the children engaged. In the picture-naming story, the experimenter read a story to the child. The target words were replaced by pictures, which the child was prompted to name. Afterwards, a speech perception task was administered for a different sub-project. The picture-naming game followed, in which a hand puppet elicited the child's speech from picture cards. When a child produced a target word more than once, every production entered the analysis. The bilinguals were tested by native speakers in two sessions that 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. Throughout the session, children were rewarded with stickers. At the end of each session, they were compensated with €10 or a book.

### Recordings and VOT measurements

Recordings were made with an Olympus Linear PCM Recorder LS-10 with uncompressed 24bit/96kHz



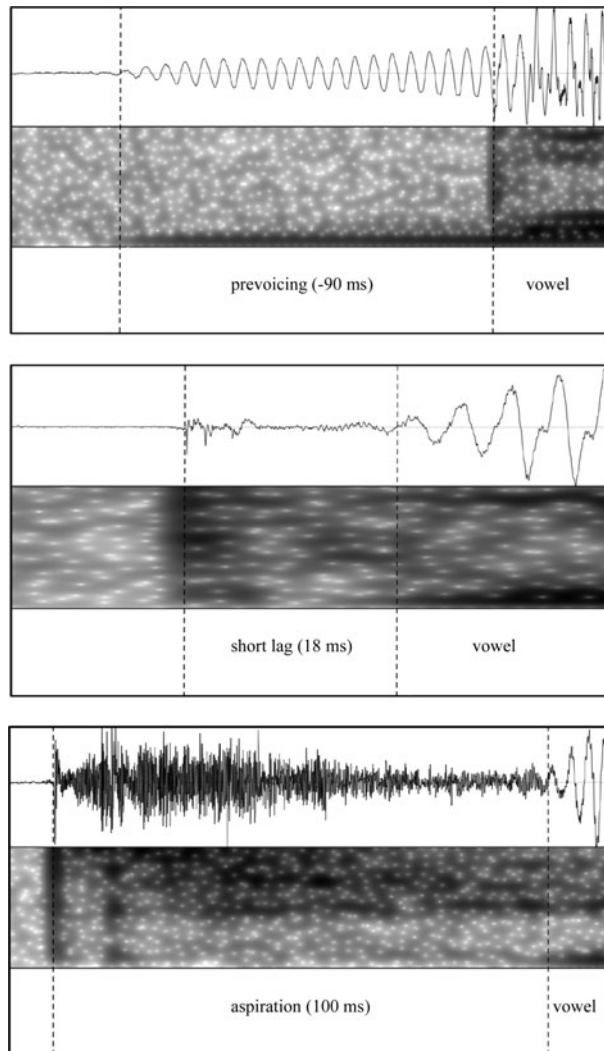


Figure 2. Acoustic landmarks from top to bottom: prevoicing, short lag, and aspiration.

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, published online 17 October, 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.

### Statistical analyses

Mixed effects models were performed in R (R Core Team, 2013). An alpha 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 included in all analyses, and Percent of Exposure to German<sup>5</sup> 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. Due to space limitations, we do not report simple effects of these IVs.

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

<sup>5</sup> 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 effects & intercept	Random slopes
Monolingual Dutch vs. monolingual German	'voiceless'	Language		Child	Task
		Age	Language*Age		PoA-LC
		Task	Language*Task		PoA-CD
		PoA-LC	Language*PoA-LC		WordLength
	'voiced'	PoA-CD	Language*PoA-CD	Item	Task
		WordLength	Language*WordLength		
		Language		Child	Task
		Age	Language*Age		PoA
Bilingual Dutch vs. bilingual German	'voiceless'	Language		Child	Language
		Age	Language*Age		Task
		%ExposureGerman	Language*%ExposureGerman		PoA-LC
		Task	Language*Task		PoA-CD
	'voiced'	PoA-LC	Language*PoA-LC		WordLength
		PoA-CD	Language*PoA-CD	Item	Task
		WordLength	Language*WordLength		
		Language		Child	Language
Bilinguals vs. monolinguals	'voiceless'	LangBackgr.		Child	Task
		Age	LangBackgr.*Age		PoA-LC
		Task	LangBackgr.*Task		PoA-CD
		PoA-LC	LangBackgr.*PoA-LC		WordLength
	'voiced'	PoA-CD	LangBackgr.*PoA-CD	Item	LangBackgr.
		WordLength	LangBackgr.*WordLength		Task
		LangBackgr.		Child	Task
		Age	LangBackgr.*Age		PoA
	Task	LangBackgr.*Task	Item	LangBackgr.	
	PoA	LangBackgr.*PoA		Task	

LangBackgr. = Language Background  
 PoA-LC = Place of Articulation: Labial vs. Coronal  
 PoA-CD = Place of Articulation: Coronal vs. Dorsal

information on post-hoc analyses is provided in Appendix S4 in the online supplementary materials (Supplementary Material).

**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 and Figure 3).

For 'voiced' plosives, monolingual Dutch children produced the highest and German monolingual children

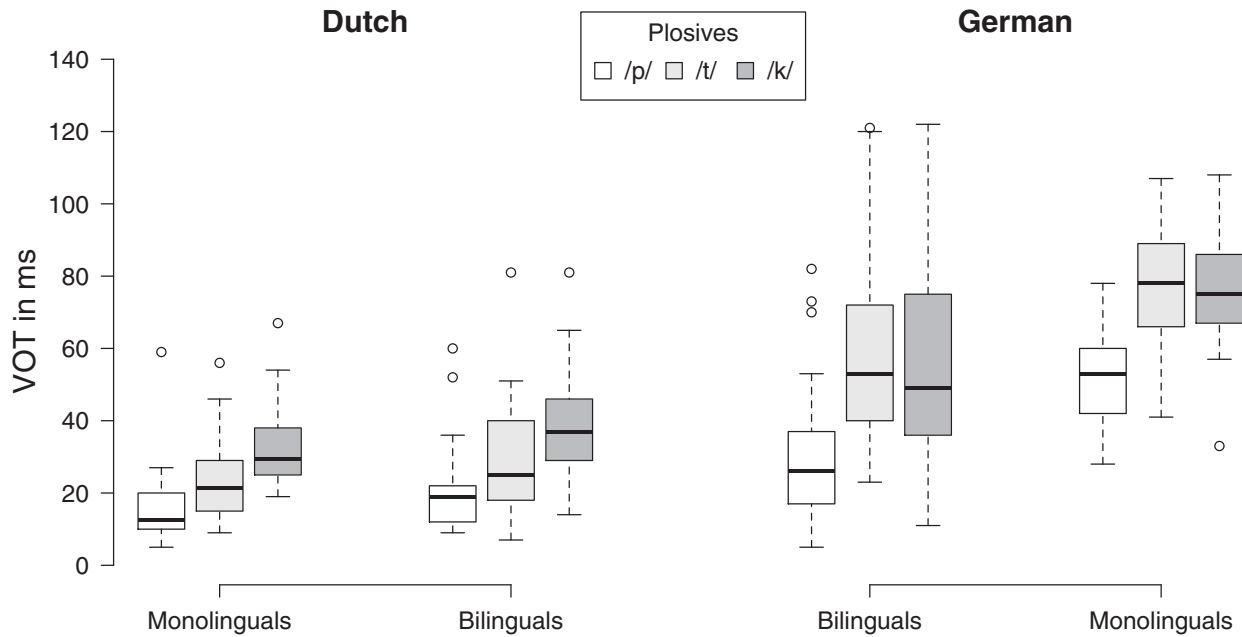


Figure 3. Voiceless plosives: VOT by language and language background over children.

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

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

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 and three bilingual children fell within the adult-like 75–100% range of prevoicing.

The devoiced ‘voiced’ plosives fell on average within the short lag VOT range. All groups produced devoiced /b/ with VOT around 10 ms. For devoiced /d/, the Dutch monolinguals and the bilinguals in both languages produced VOT 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 (cf. Tables 4 and 6). Figure 5 shows the distribution of VOT across all ‘voiced’ plosives by group and language.

Four sets of mixed effects regression 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 ( $\beta = 28.96$ ,  $SE = 2.95$ ,  $t = 9.82$ ,  $p < .001$ ). Interactions between Language and Place of Articulation (labial vs. coronal;  $\beta = -7.28$ ,  $SE = 3.46$ ,  $t = -2.10$ ,  $p = .036$ ) as well as Language and Word Length ( $\beta = 4.33$ ,  $SE = 1.36$ ,  $t = 3.18$ ,  $p = .002$ ) indicated that the German monolingual children produced shorter VOT in labial /p/ than in coronal /t/ ( $\beta = -21.34$ ,  $SE = 5.89$ ,  $t = -3.63$ ,  $p < .001$ ) and longer VOT in monosyllabic than in disyllabic words ( $\beta = 8.79$ ,  $SE = 2.37$ ,  $t = 3.71$ ,  $p < .001$ ), but neither effect was observed

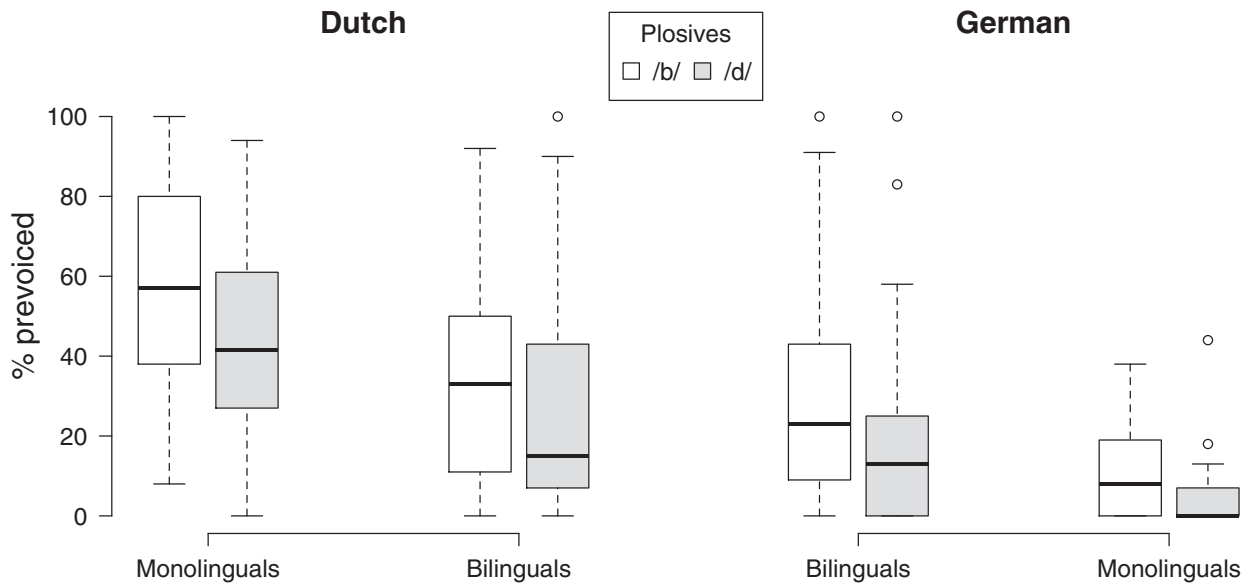


Figure 4. Voiced plosives: percent prevoiced by language and language background over children.

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

	Dutch		Bilinguals		German		Monolinguals	
	prevoiced	devoiced	prevoiced	devoiced	prevoiced	devoiced	prevoiced	devoiced
/b/								
<b>M %</b>	<b>56</b>	<b>44</b>	<b>33</b>	<b>67</b>	<b>30</b>	<b>70</b>	<b>12</b>	<b>88</b>
<i>SD</i>	27	27	27	27	27	27	13	13
Total #	177	139	112	228	112	258	48	333
/d/								
<b>M %</b>	<b>43</b>	<b>57</b>	<b>28</b>	<b>72</b>	<b>21</b>	<b>79</b>	<b>4</b>	<b>96</b>
<i>SD</i>	25	25	29	29	26	26	8	8
Total #	155	194	97	247	83	329	15	414
Range prevoiced	9–96%		0–95%		0–100%		0–41%	
Total# /b/+/d/	332	333	209	475	195	587	63	747

in the monolingual Dutch children ( $\beta = -6.33, SE = 3.60, t = -1.76, p = .079$  and  $\beta = 0.09, SE = 1.31, t = .06, p > .250$ , respectively). Monolingual Dutch children produced a higher percentage of ‘voiced’ plosives with prevoicing than monolingual German children ( $\beta = 1.56, SE = 0.19, z = 8.07, p < .001$ ). An interaction between Language and Place of Articulation ( $\beta = -0.21, SE = 0.11, z = -2.03, p = .042$ ) indicated that both groups prevoiced labial /b/ more frequently than coronal /d/, but the magnitude of the effect was larger in the German monolinguals ( $\beta = -1.06, SE = 0.23, z = -4.52, p < .001$ ) than in the Dutch monolinguals ( $\beta = -0.32, SE = 0.12, z = -2.80, p = .005$ ). 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 ( $\beta = 14.43, SE = 3.54, t = 4.08, p < .001$ ). An interaction between Language and Percent of Exposure to German ( $\beta = 0.26, SE = 0.09, t = 2.85, p = .004$ ) revealed that more exposure to German is associated with longer, and therefore more

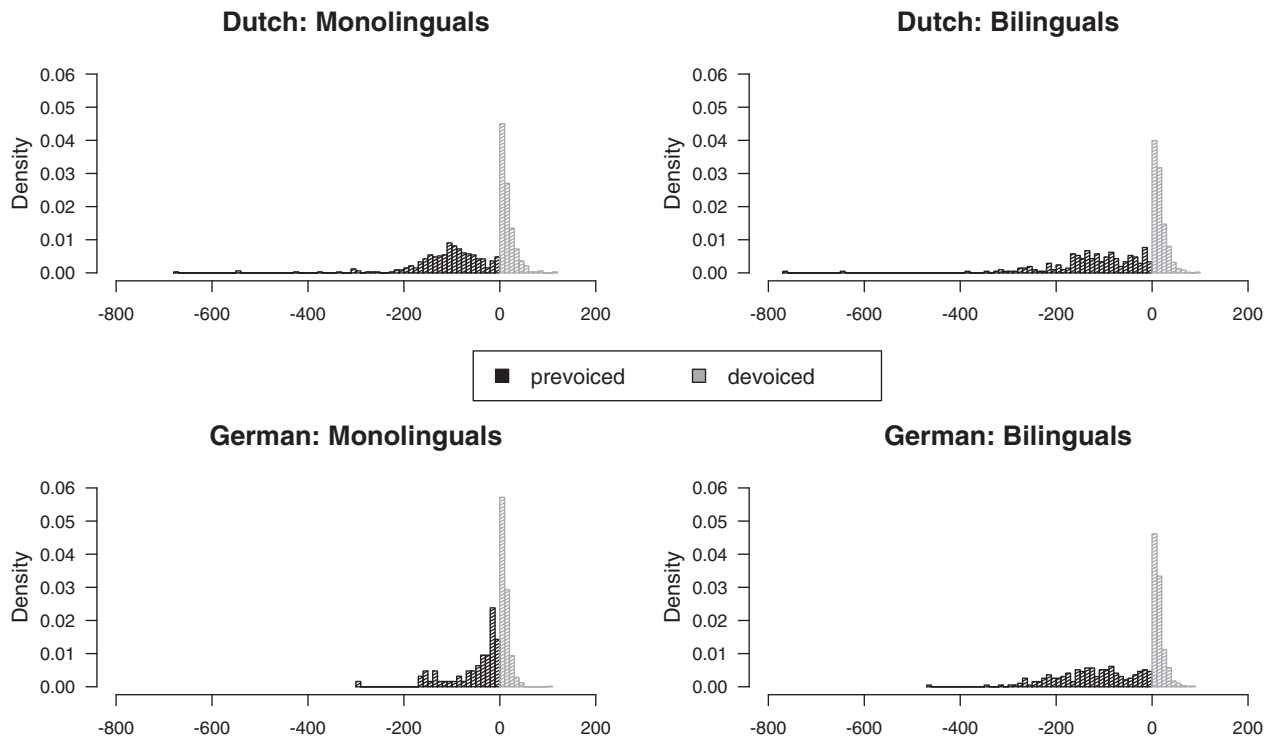


Figure 5. VOT distribution of voiced plosives by language and language background.

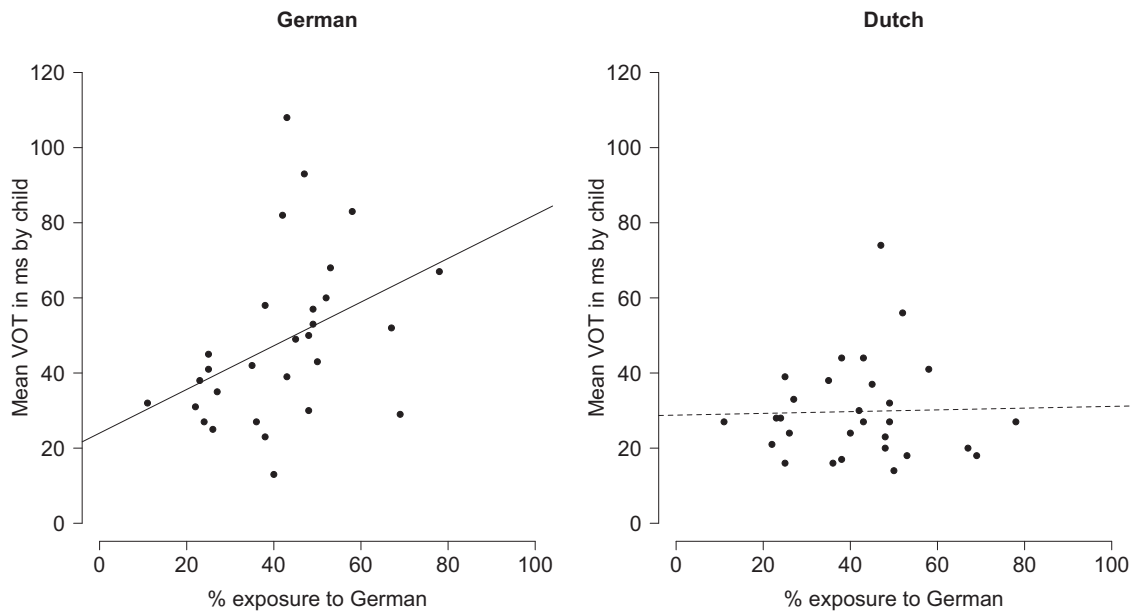


Figure 6. VOT of voiceless plosives and exposure to German over children.

target-like VOT in German ( $\beta = 0.52$ ,  $SE = 0.24$ ,  $t = 2.17$ ,  $p = .030$ ), while it had no detectable effect on the bilinguals' Dutch VOT ( $\beta = 0.15$ ,  $SE = 0.14$ ,  $t = 1.09$ ,  $p > .250$ ) as visualized in Figure 6. Similarly, an interaction between Language and Task ( $\beta = 1.07$ ,  $SE = 0.49$ ,  $t = 2.21$ ,  $p = .027$ ) indicated that the bilinguals produced longer VOT in the game task than in the story task in

Dutch ( $\beta = -3.14$ ,  $SE = 0.92$ ,  $t = -3.42$ ,  $p < .001$ ), but not in German ( $\beta = -1.11$ ,  $SE = 1.25$ ,  $t = -0.89$ ,  $p > .250$ ). The percentage of 'voiced' plosives produced with prevoicing was similar in the bilinguals' Dutch and German ( $\beta = 0.26$ ,  $SE = 0.17$ ,  $z = 1.5$ ,  $p = .134$ ), and it was not significantly affected by Percent of Exposure to German ( $\beta = 0.02$ ,  $SE = 0.02$ ,  $z = 1.22$ ,  $p = .223$ ).

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
/b/				
<b>M</b>	<b>11</b>	<b>13</b>	<b>11</b>	<b>10</b>
<i>SD</i>	7	5	6	4
Total #	139	229	258	333
/d/				
<b>M</b>	<b>20</b>	<b>21</b>	<b>18</b>	<b>13</b>
<i>SD</i>	10	7	5	4
Total #	194	247	329	414
Group Total#	333	476	587	747

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 ( $\beta = 2.86$ ,  $SE = 1.92$ ,  $t = 1.50$ ,  $p = .134$ ). However, the bilinguals produced a lower percentage of prevoiced 'voiced' plosives in Dutch than their monolingual peers ( $\beta = 0.51$ ,  $SE = 0.21$ ,  $z = 2.40$ ,  $p = .016$ ).

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

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':  $\beta = 0.17$ ,  $SE = 0.15$ ,  $t = 1.2$ ,  $p = .230$  & 'voiced':  $\beta = -0.02$ ,  $SE = 0.02$ ,  $z = -1.14$ ,  $p > .250$ ; bilingual Dutch vs. bilingual German: 'voiceless':  $\beta = -0.23$ ,  $SE = 0.29$ ,  $t = -0.80$ ,  $p > .250$  & 'voiced':  $\beta = -0.02$ ,  $SE = 0.03$ ,  $z = -0.72$ ,  $p > .250$ ; bilingual Dutch vs. monolingual Dutch: 'voiceless':  $\beta = 0.09$ ,  $SE = 0.14$ ,  $t = 0.62$ ,  $p > .250$  & 'voiced':  $\beta = -0.03$ ,

$SE = 0.02$ ,  $z = -1.57$ ,  $p = .116$ ; bilingual German vs. monolingual German: 'voiceless':  $\beta = -0.15$ ,  $SE = 0.25$ ,  $t = -0.58$ ,  $p > .250$  & 'voiced':  $\beta = -0.03$ ,  $SE = 0.02$ ,  $z = -1.05$ ,  $p > .250$ ).

## 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 & Abramson, 1964; Neuhauser, 2011). Equivalent to Dutch and German monolingual children, the bilinguals produced longer VOT in German than in Dutch, suggesting 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

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

		Monolingual Dutch vs. monolingual German	<b>RQ 1</b> <b>Bilingual Dutch</b> <b>vs. bilingual</b> <b>German</b>	<b>RQ 2</b> <b>Bilingual Dutch</b> <b>vs. monolingual</b> <b>Dutch</b>	<b>Bilingual German</b> <b>vs. monolingual</b> <b>German</b>
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.	–	–
<b>RQ 3: Age</b>	'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.  $p > .05$

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, published online October 17, 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, published online 3 May, 2017).

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., published online 3 May, 2017; 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; Khatlab, 2000; Macken & Barton, 1980b; MacLeod, 2016). 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 falls 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., published online 3 May, 2017).

However, the present data is 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; cf. 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., published online 3 May, 2017). The bilinguals' exposure



to Dutch leads to more exposure to prevoicing, and 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.

### Conclusion

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.

### Supplementary Material

For supplementary material accompanying this paper, visit <http://dx.doi.org/10.1017/S1366728917000116>

### 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*. <http://www.praat.org/> (retrieved December 28, 2014).
- 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, 113–133. <http://www.fon.hum.uva.nl/archive/1976/Proc4Deighton.pdf> (retrieved December 3, 2015).
- 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. (2003). 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, 169–81.
- 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, Volume 8*, pp. 45–70. [http://www.leeds.ac.uk/arts/info/125154/leeds\\_working\\_papers\\_in\\_linguistics\\_and\\_phonetics/1939/volume\\_8\\_2000](http://www.leeds.ac.uk/arts/info/125154/leeds_working_papers_in_linguistics_and_phonetics/1939/volume_8_2000) (retrieved December 3, 2015).
- 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. <http://ismbs.eu/data/documents/Proceedings-ISMBMS-2015.pdf> (retrieved February 10, 2016).
- 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, Volume 8*, pp. 95–122. [http://www.leeds.ac.uk/arts/info/125154/leeds\\_working\\_papers\\_in\\_linguistics\\_and\\_phonetics/1939/volume\\_8\\_2000](http://www.leeds.ac.uk/arts/info/125154/leeds_working_papers_in_linguistics_and_phonetics/1939/volume_8_2000) (retrieved December 3, 2015).
- 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*, doi: 10.1177/1367006916654355. Published online by Sage Publishing, June 22, 2016.
- 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. (2016). 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*, doi: 10.1177/1367006916672590. Published online by Sage Publishing, October 17, 2016.
- 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, Hong Kong*, pp. 1462–1465. <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/index.htm> (retrieved December 3, 2015).
- 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. <http://www.R-project.org/> (retrieved March 30, 2014).
- 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 Phonemrealisationen*. 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*, doi: 10.1177/0267658317704261. Published online by Sage Publishing, May 3, 2017.
- Szagan, G., Stumper, B., & Schramm, S. A. (2009). *FRAKIS - Fragebogen zur frühkindlichen Sprachentwicklung*. *PSYNDEX Tests Info*. Frankfurt am Main: Pearson.
- Tamburelli, M., Sanoudaki, E., Jones, G., & 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: 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., & Koenigsnecht, 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.