Feature generalization in Dutch–German bilingual and monolingual children's speech production

Stoehr, A.; Benders, T.; van Hell, J.G.; Fikkert, P.

DOI
10.1177/01427237211058937

Publication date
2022

Document Version
Final published version

Published in
First Language

License
CC BY

Link to publication

Citation for published version (APA):
Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2022). Feature generalization in Dutch–German bilingual and monolingual children's speech production. *First Language*, *42*(1), 101-123. https://doi.org/10.1177/01427237211058937

Article

FIRST LANGUAGE

First Language
2022, Vol. 42(1) 101–123

DOI: 10.1177/01427237211058937
journals.sagepub.com/home/fla
SAGE

Feature generalization in Dutch–German bilingual and monolingual children's speech production

Antje Stoehr
Radboud University, The Netherlands; International Max Planck Research School for Language Sciences, The Netherlands; Basque Center on Cognition, Brain and Language, Spain

Titia Benders
Macquarie University, Australia

Janet G. van Hell
The Pennsylvania State University, USA

Paula Fikkert
Radboud University, The Netherlands

Abstract
Dutch and German employ voicing contrasts, but Dutch lacks the 'voiced' dorsal plosive /g/. We exploited this accidental phonological gap, measuring the presence of prevoicing and voice onset time durations during speech production to determine (1) whether preliterate bilingual Dutch–German and monolingual Dutch-speaking children aged 3;6–6;0 years generalized voicing to /g/ in Dutch; and (2) whether there was evidence for featural cross-linguistic influence from Dutch to German in bilingual children, testing monolingual German-speaking children as controls. Bilingual and monolingual children's production of /g/ provided partial evidence for feature generalization: in Dutch, both bilingual and monolingual children either recombined Dutch voicing and place features to produce /g/, suggesting feature generalization, or

Corresponding author:
Antje Stoehr, Basque Center on Cognition, Brain and Language, Paseo Mikeletegi 69, Donostia–San Sebastián 20009, Spain.
Email: a.stoehr@bcbl.eu

resorted to producing familiar /k/, suggesting segment-level adaptation within their Dutch phonological system. In German, bilingual children's production of /g/ was influenced by Dutch although the Dutch phoneme inventory lacks /g/. This suggests that not only segments but also voicing features can exert cross-linguistic influence. Taken together, phonological features appear to play a crucial role in aspects of bilingual and monolingual children's speech production.

Keywords
Speech production, phonological features, simultaneous bilinguals, cross-linguistic influence, preschoolers, Dutch, German

When children acquire a phonological system, they initially distinguish broad contrasts in production, such as the contrast between labial and non-labial plosives. Subsequently, they acquire narrower contrasts, such as the voicing contrast between 'voiced'[1] plosives /bdg/ and 'voiceless' plosives /ptk/ (Brown & Matthews, 1997; Clements, 2003, 2005; Dresher, 2004; Jakobson, 1941; Rice & Avery, 1995). In phonological theory, contrasts are captured by phonological features (Chomsky & Halle, 1968; Clements, 1985; Clements & Hume, 1995; Durand & Laks, 2002; Halle, 1964; Kiparsky, 1995; Mielke, 2008). For example, the voicing contrast can be captured by the phonological feature [voice].[2] Despite the importance of phonological features in phonological theory and their acknowledged role in adult speech perception and production, their role remains controversial in children's speech perception (e.g. Cristià et al., 2011; Mani & Plunkett, 2007; Pater et al., 2004; Stager & Werker, 1997; Swingley & Aslin, 2000; Werker et al., 2002) and production (Menn & Vihman, 2011). The present study provides evidence that phonological features influence *production* patterns in both monolingual and bilingual preschoolers.

Accidental phonological gaps provide a unique opportunity to investigate whether children's speech production can be based on generalizable phonological features, for example, [voice], or is only related to whole segments, for example, /g/. An accidental phonological gap refers to the absence of a specific phonemic contrast in part of a language's phonological system as a result, for example, of historical changes (e.g. Iverson & Salmons, 2005). For instance, Dutch labial and coronal plosives exhibit voicing contrasts ('voiceless' /p/ and /t/ vs 'voiced' /b/ and /d/), but the 'voiceless' dorsal plosive /k/ lacks a 'voiced' counterpart /g/, leading to an accidental phonological gap in the Dutch phoneme inventory. In fact, several languages lack /g/, arguably because /g/ is relatively difficult to produce (Gussenhoven & Jacobs, 2017; Ohala, 1981, 2010).

The present study exploits the accidental phonological gap /g/ in Dutch to investigate whether feature generalization influences children's speech production. Specifically, this study addresses feature generalization both *within* the Dutch phonological system of monolingual and bilingual children and *between* bilingual children's Dutch and German phonological systems. In the following sections, we first outline the current understanding of feature generalization within and between languages and describe the voicing systems of Dutch and German. We then present our current study.

Within-language feature generalization

Speech perception experiments have shown that phonological features guide adult and even infant native language processing (among many others, e.g. Chládková et al., 2017; Finley & Badecker, 2009; Hestvik & Durvasula, 2016; Kraljic & Samuel, 2006; Lahiri & Reetz, 2002, 2010). For instance, adult native speakers of English generalize back-vowel harmony in an artificial grammar to new vowels only presented during the test (Finley & Badecker, 2009). Moreover, after exposure to ambiguous /d/-/t/ tokens in place of either target /d/ (e.g. *crocodile*) or target /t/ (e.g. *cafeteria*), English native speakers shift their perceptual boundary between /d/ and /t/, and also between /b/ and /p/, suggesting that perceptual learning occurs on the feature-level (Kraljic & Samuel, 2006). Feature generalization does not seem to be restricted to adult language processing. Even in the early stages of language acquisition, infants appear to generalize phonetic or acoustic features (for an overview, see Cristià et al., 2011). For example, when trained on the Hindi voicing contrast between prevoicing and short-lag voice onset time (VOT) at the dental place of articulation, 8-month-old English-acquiring infants generalized this contrast to the dorsal place of articulation during the test (Maye et al., 2008). Taken together, these studies suggest that the speech perception of (to-be) native English speakers is guided by feature generalization.

While evidence is sparse, two studies suggest that feature generalization also plays a role in adult speech production. Adult native speakers of English exposed to extreme VOT durations for /p/ in an imitation task produced /p/ as well as /k/ with lengthened VOT (Nielsen, 2011). Similarly, English-speaking learners of Spanish trained on Spanish short-lag VOT for a ‘voiceless’ plosive at one place of articulation generalized this VOT shortening to ‘voiceless’ plosives at the other two places of articulation (Olson, 2019). These two studies suggest that phonetic convergence can generalize across sounds that share a feature.

More naturalistic evidence for the role of features in speech production comes from studies on Dutch-speaking adults’ production of the non-native segment /g/. Although the segment /g/ corresponds to an accidental phonological gap in the Dutch plosive inventory of /bd ptk/, it still occurs in a number of loanwords in Dutch (Hamann & de Jonge, 2015; Van Bezooijen & Gerritsen, 1994). Dutch-speaking adults predominantly produce word-initial /g/ as [g], but [k] and [χ/x] (depending on the dialect) have also been observed. Production of /g/ as [g] suggests that Dutch speakers generalize the feature [voice] to the dorsal place of articulation. However, most Dutch adults are already familiar with /g/ from English or another second language (L2; Ytsma, 2000). It therefore remains unclear whether their production of [g] is based on feature generalization within their native (L1) phonology or reflects segmental borrowing from L2 phonology. The production of /g/ as [k] instead suggests that Dutch native speakers adapt non-native /g/ to the perceptually closest native segmental category. The production of /g/ as [χ/x] is likely related to Dutch orthography, in which /χ/x/ corresponds to the grapheme <g>. To evaluate whether feature generalization within the L1 phonology is at play in /g/ production by Dutch native speakers, it is necessary to get insight from native speakers who are neither literate nor exposed to an L2. Preschoolers raised in monolingual families fulfill both criteria.

Between-language feature generalization

If speech production is indeed guided by features, this should be the case for monolinguals and bilinguals alike. It is well established that the phonologies of most bilingual children interact to some extent, a phenomenon known as cross-linguistic influence (Deuchar & Clark, 1996; Fabiano-Smith & Bunta, 2012; Johnson & Wilson, 2002; Kehoe et al., 2004; Khattab, 2000; Stoehr et al., 2018). Yet, it remains unclear whether cross-linguistic influence between a bilingual's two languages operates only on specific phonological segments or also on phonological features.

The only support for cross-linguistic influence of features comes from a study on adult L2 learners of English, whose native Arabic dialect lacks the segment /p/ (Flege & Port, 1981). These learners produced English /p/ as unaspirated, which could be evidence for feature-level cross-linguistic influence of their L1 Arabic /t/ and /k/ on the novel L2 segment /p/. Unaspirated /p/, however, is also articulatory simpler than the aspirated target (Kewley-Port & Preston, 1974), and the unaspirated realization of /p/ could thus also be explained by general articulatory restrictions rather than feature-level cross-linguistic influence. To reliably test whether cross-linguistic influence operates on features, it is thus necessary to investigate language pairs in which cross-linguistic influence between features would result in articulatory more complex production. The segment /g/, which is part of the German but not the Dutch phoneme inventory offers a unique test case, as feature-level cross-linguistic influence of Dutch on German would result in the articulatorily more complex, prevoiced production in German, as explained in the following section.

Voicing in Dutch and German

The 'voiced' dorsal plosive /g/ is native to German but non-native to Dutch. However, it is composed of features that form part of Dutch phonology: /g/ carries the feature [plosive], shared with Dutch /b/, /d/, /p/, /t/ and /k/; the feature [dorsal], shared with Dutch /k/ (among others); and the feature [voice], shared with Dutch /b/ and /d/ (among others).

Importantly, Dutch and German differ in the phonetic implementation of the voicing contrast in terms of VOT, which is the primary cue to voicing in word-initial plosives (Jessen, 1998; Lisker & Abramson, 1964). VOT is the time interval between the release of a plosive (e.g. release of the closure between the back of the tongue and the velum for /g/ and /k/) and the onset of voicing, that is, vocal fold vibration. VOT is measured in milliseconds and falls within three major categories (Keating, 1984; Lisker & Abramson, 1964): prevoicing (voicing starts prior to burst release), short-lag (voicing starts around 0–30 ms after burst release), and aspiration (voicing starts later than 30 ms after burst release). Dutch has short-lag 'voiceless' plosives and prevoiced 'voiced' plosives. In German, 'voiceless' plosives are aspirated, and 'voiced' plosives typically have short-lag VOT, although some instances of prevoicing have been observed in both monolingual adults' (Stoehr et al., 2017) and monolingual children's (Stoehr et al., 2018) production of German 'voiced' plosives.

In the early stages of speech production, children's plosive production generally falls within the short-lag VOT range irrespective of their L1: German monolingual children tend to omit aspiration and produce ‘voiceless’ plosives with short-lag VOT (Kager et al., 2007), while Dutch monolingual children often omit prevoicing and produce ‘voiced’ plosives with short-lag VOT. The tendency to omit prevoicing can even be observed in Dutch production by Dutch monolingual children and Dutch–German bilingual children up to the age of 6 years (Stoehr et al., 2018). Dutch monolingual children aged between 3;6 and 6;0 years devoice approximately half of their word-initial /b/ and /d/ productions. Age-matched Dutch–German bilingual children follow the same pattern in Dutch, omitting prevoicing in two-thirds of their /b/ and /d/ production. Importantly, these monolingual and bilingual Dutch-speaking children produce short-lag /b/ and /d/ with shorter VOT than /p/ and /t/, thus maintaining the voicing contrast even though both ‘voiced’ and ‘voiceless’ plosives are produced within the short-lag range. At age 3;6–6;0, German monolingual children seem to have acquired native-like aspiration in production. Age-matched Dutch–German bilingual children similarly produce aspirated ‘voiceless’ plosives in German, although the duration of aspiration is shorter than that of their monolingual peers (Stoehr et al., 2018).

The present study

The present study takes an acoustic-phonetic approach to investigate the production of /g/ by monolingual Dutch-speaking children (hereafter, Dutch monolinguals), bilingual Dutch–German speaking children (hereafter, bilinguals) and monolingual German-speaking controls (hereafter, German monolinguals). All children were preliterate and had no exposure to additional languages. For this reason, production of the non-Dutch segment /g/ by Dutch monolinguals could only be explained by processes within their native phonology, while bilinguals’ production of the non-Dutch segment /g/ could be explained either by processes within their Dutch phonology or by the influence of their German phonology.

Based on Dutch adults’ production of /g/ in loanwords from English incorporated into Dutch (Hamann & de Jonge, 2015; Van Bezooijen & Gerritsen, 1994), we can hypothesize two possible outcomes for preliterate Dutch monolingual children’s production of /g/, as summarized in Table 1: generalization of the feature [voice] to the dorsal place of articulation (Table 1, pattern A), or production of the perceptually similar native-segment /k/ (Table 1, pattern B). Bilinguals who acquire Dutch (lacking /g/) and German (encompassing /g/) might produce /g/ in a Dutch context in the same two ways as Dutch monolinguals. Thus, a possible outcome for both monolinguals and bilinguals would be prevoicing a proportion of /g/ productions as the Dutch feature [voice] is generalized from /b/ and /d/ to the dorsal place of articulation (Table 1, pattern A). Nevertheless, as prevoicing is more difficult to produce in places of articulation further back in the mouth, the proportion of prevoiced tokens is expected to be lower for /g/ than for /b/ and /d/ (Cho & Ladefoged, 1999; Ohala & Riordan, 1979). The other possible outcome for both monolinguals and bilinguals would be resorting to /k/ produced with short-lag VOT of approximately 30 ms (for adults: Lisker & Abramson, 1964; Stoehr et al., 2017; for children: Stoehr et al., 2018) as a result of adaptation (Table 1, pattern B). In this context, adaptation refers to a process in which non-native segments are mapped onto L1 phonemes (Silverman, 1992). Both scenarios would reflect processes within the children’s

Table 1. Possible production patterns for /g/ by group and language.

	Pattern	Possible production patterns for /g/ in Dutch			Possible production patterns for /g/ in German	
		A: Feature generalization within Dutch	B: Segmental adaptation within Dutch	C: Segmental adaptation from German	D: Featural CLI from Dutch	E: No featural influence from Dutch
Dutch	Monolinguals	Some /g/ as prevoiced, in line with /b/ and /d/		—	—	—
		negative VOT		—	—	—
		Some /g/ as short-lag	/g/ as /k/	—	—	—
		VOT ~ 0 ms	*VOT ~ 30 ms*	—	—	—
	Bilinguals	Some /g/ as prevoiced, in line with /b/ and /d/		Very few /g/ as prevoiced	—	—
		negative VOT		*negative VOT*	—	—
		Some /g/ as short-lag	/g/ as /k/	German /g/	—	—
		VOT ~ 0 ms	*VOT ~ 30 ms*	*VOT ~ 15 ms*	—	—
German	Bilinguals	—	—	—	Some /g/ as prevoiced, in line with /b/ and /d/	Very few /g/ as prevoiced
		—	—	—	*negative VOT*	*negative VOT*
		—	—	—	Some /g/ as short-lag	Monolingual-like /g/
		—	—	—	*VOT ~ 0 ms*	*VOT ~ 15 ms*

Dutch phonological system. Alternatively, Dutch–German bilingual children may use the German segment /g/ within a Dutch context (Table 1, pattern C), which would primarily result in short-lag VOT of approximately 15 ms (Braunschweiler, 1997 for adults; Kehoe et al., 2004 for German-speaking children aged 1;9–2;6), in addition to a very low proportion of prevoiced /g/ tokens (Stoehr et al., 2018).

Importantly, Dutch–German bilinguals' production of /g/ in German cannot be subject to segmental influence from Dutch, in which /g/ is non-existent. This allows us to address cross-linguistic influence of features in bilingual speech production: if there is cross-linguistic influence of Dutch on German, resulting in prevoiced production of /g/ in German despite the absence of /g/ in Dutch (Table 1, pattern D), this could be taken as evidence of cross-linguistic influence at the feature-level. Here, cross-linguistic influence at the feature-level would result in bilingual children producing German /g/ with prevoicing more frequently than monolinguals, that is, they would use the more articulatorily complex alternative. This would rule out any kind of 'default' production (Kewley-Port & Preston, 1974; Macken & Barton, 1980), which could have explained previous findings on Arabic speakers' production of unaspirated /p/ in English (Flege & Port, 1981). If, on the contrary, bilinguals produce German /g/ like monolinguals (Table 1, pattern E), it would suggest no featural influence of Dutch on German.

Taken together, this article thus addresses the overarching research question of whether the production of /g/ by monolingual Dutch children, bilingual children speaking Dutch, and bilingual children speaking German can be understood as the result of feature-level influence or segmental adaptation. We will address this general question in a series of three analyses: the first analysis targets monolinguals' and bilinguals' production of word-initial /g/ in Dutch, the second analysis is concerned with bilinguals' production of /g/ in German compared with German monolingual controls, and the third analysis compares the bilinguals' production of /g/ in Dutch and German. Each analysis addresses one or more of the outlined hypotheses and the combined results will reveal whether children's production of /g/ is better explained by feature-level influence or segmental adaptation.

1. Dutch production of monolingual and bilingual children
    a. Do Dutch monolingual children and bilingual children alike generalize the feature [voice] to the dorsal place of articulation, resulting in instances of prevoiced /g/ in Dutch (Table 1, pattern A)?
    b. When producing /g/ *without* prevoicing, do Dutch monolingual and bilingual children produce these devoiced instances of /g/ in line with their Dutch segment /k/ (Table 1, pattern B)?
2. German production of bilingual children and monolingual controls
    a. Do bilingual children prevoice /g/ more frequently in German than German monolinguals, providing evidence for feature-level cross-linguistic influence from Dutch, which lacks /g/ (Table 1, pattern D) or do they produce /g/ like monolinguals (Table 1, pattern E)?
    b. When producing /g/ in German *without* prevoicing, do bilinguals produce shorter VOT than monolinguals, providing evidence for feature-level

cross-linguistic influence from Dutch, which lacks /ɡ/ (Table 1, pattern D) or do they produce /ɡ/ like monolinguals (Table 1, pattern E)?

3. Bilingual children in Dutch and German
   a. When producing /ɡ/ *with* prevoicing, do bilingual children employ their German /ɡ/ segment in Dutch (Table 1, pattern C) or do they generalize their Dutch voicing feature to /ɡ/ in production (Table 1, pattern A)?
   b. When producing /ɡ/ *without* prevoicing, do bilingual children employ their German /ɡ/ segment in Dutch (Table 1, pattern C) or do they generalize their Dutch voicing feature to /ɡ/ in production (Table 1, pattern A)?

Given previous findings that support the existence of features in adult speech perception and production (Chládková et al., 2017; Finley & Badecker, 2009; Hestvik & Durvasula, 2016; Kraljic & Samuel, 2006; Lahiri & Reetz, 2002, 2010; Nielsen, 2011; Olson, 2019) and infant speech perception (Maye et al., 2008), we hypothesized that production of the segment /ɡ/ in Dutch by monolingual Dutch and Dutch–German bilingual children would be guided by features (Table 1, pattern A). Moreover, we hypothesized that cross-linguistic influence from Dutch to German would operate on features and result in instances of prevoiced dorsal plosives in the bilingual children's German (Table 1, pattern D).

Method

Participants

Eighty-seven children aged 3;6–6;0 years participated in this study: 30 Dutch monolinguals (M_{age} = 4;9, range 3;6–6;0; 17 female), 28 Dutch–German simultaneous bilinguals (M_{age} = 4;7, range 3;7–5;11; 13 female), and 29 German monolinguals (M_{age} = 4;8, range 3;6–6;0; 20 female). The groups did not differ significantly in age, $F(2, 84) = 0.38$, $p > .250$. The same children were tested in Stoehr et al. (2018).

Based on parental report, all children were typically developing and none had been exposed to an additional language or foreign-accented speakers. All bilinguals were born and raised in a Dutch-dominant environment in the Netherlands where they went to Dutch-speaking daycare centers. They had acquired German as a heritage language from birth spoken as a native language in the home by either one ($N = 25$) or both ($N = 3$) parents. Among the three bilingual children born to two German-speaking parents, two were exposed to Dutch from birth through other caretakers, and the third was first exposed to Dutch at 6 months of age at a daycare center. Only bilinguals able to communicate in Dutch and German were invited to participate. Parents provided detailed information on their bilingual children's language exposure based on the Bilingual Language Experience Calculator (Unsworth, 2013). Bilingual children were exposed more to Dutch ($M = 60\%$, range 31%–89%, $SD = 14$) than to German ($M = 40\%$, range 11%–69%, $SD = 14$) at the time of testing, $t(27) = 3.615$, $p = .001$. Based on parental proficiency ratings from 0 (virtually no fluency; almost no understanding) to 5 (native fluency, native understanding), bilingual children had better productive proficiency in Dutch ($M = 4.7$, range 2–5,

$SD = 0.7$) than in German ($M = 3.3$, range 1–5, $SD = 1.2$), $t(27) = 4.837$, $p < .001$. They also had better receptive proficiency in Dutch ($M = 4.9$, range 3–5, $SD = 0.5$) than in German ($M = 4.6$, range 3–5, $SD = 0.6$), $t(27) = 2.295$, $p = .030$. Of the 97 children initially tested, five bilinguals and five monolinguals were excluded because they did not meet the inclusion criteria ($N = 8$) or were unable to complete the task ($N = 2$). 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.

Materials and procedure

Production of word-initial 'voiced' and 'voiceless' dorsal plosives /g/ and /k/ was investigated. In addition, six tokens each of /b/, /p/, /d/, and /t/ were elicited in the same tasks for a different project (Stoehr et al., 2018). For each native segment (Dutch /k/; German /g/ and /k/), six items were selected from developmental vocabulary lists (for Dutch: Zink & Lejaegere, 2002; for German: Grimm & Doil, 2000; Szagun et al., 2009). For the Dutch non-native segment /g/, one English loanword and three names complying with Dutch phonotactics were used to elicit production (items listed in Table S1, Supplementary material). Each name referred to a character introduced to the children at the beginning of the testing session. The experimenter repeated each name 10 times (Singh et al., 2014; Woodward et al., 1994), producing the initial /g/ as prevoiced.[3] As indicated by parental report on the bilingual children, none of the children had prior knowledge of the names used in the task, but some children knew the loanword 'goal'. For this reason, we ran all analyses both including and excluding 'goal'. The pattern of results did not change when this item was excluded. The Results section, therefore, only reports those statistical analyses that included all items.

Each child was tested in a quiet room at home. Parents gave informed consent and completed a language background questionnaire. The children were asked to name all items in two different picture-naming tasks to maximize the number of tokens produced per child. In the picture-naming story, the experimenter read a story in which pictures replaced the items. When the experimenter encountered a picture in the sentence, they would, for example, point or raise their intonation to encourage the child to name the picture. In the picture-naming game, children were shown cards with characters and objects and asked to name these for a hand-puppet, who was introduced as not knowing any of the names and words. Three native speakers of Dutch (one male) administered the Dutch testing sessions and three female native speakers of German administered the German sessions. The two sessions for the bilinguals were scheduled approximately 2 weeks apart, and the session order was counterbalanced by language. Children were rewarded with stickers during the tasks and compensated with €10 or a book at the end of each session.

Recordings and VOT measurements

Recordings were made with an Olympus Linear PCM Recorder LS-10 with uncompressed 24 bit/96 kHz recording capability. The first author measured children's VOT in Praat (Boersma & Weenink, 2015), considering both waveforms and spectrograms

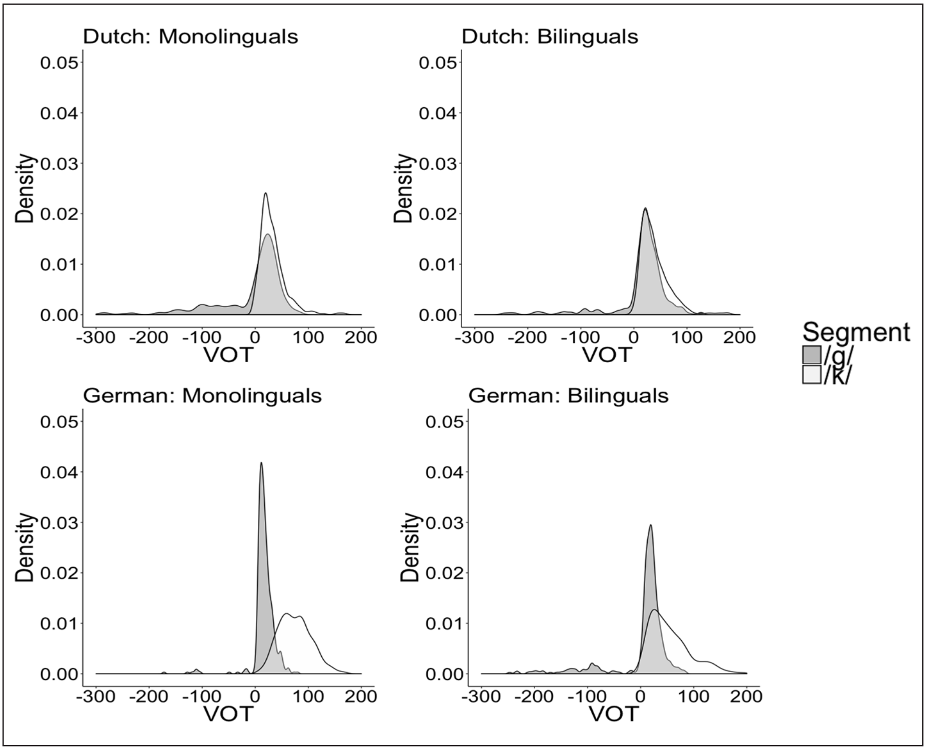

Figure 1. VOT Distribution of /g/ and /k/ in Dutch (Top) and German (Bottom) by Language Background.

viewed at 0–5000 Hz. VOT was measured as the interval between the burst onset and the onset of voicing. Burst onsets were marked at the onset of abrupt energy release. If there was more than one release burst, VOT was measured from the first fully aspirated burst (Stoel-Gammon et al., 1994). Voicing onset was defined as the first periodic component of the waveform and was measured at the preceding zero-crossing (Francis et al., 2003). Of the 3347 productions of /g/ (1606) and /k/ (1741), 2962 (1414 /g/; 1548 /k/) were analyzed (88.5%; 88% of /g/; 88.9% of /k/).[4] The remaining 385 productions (11.5%) of /g/ (192; 12%) and /k/ (193; 11.1%) were excluded from the analyses because they could not be unambiguously measured due to, for example, coarticulation, sound overlap, or whispering.

Results

Descriptive statistics

All three groups of children produced some prevoiced instances of /g/, but the majority of /g/ productions were produced without prevoicing (Figure 1). In Dutch, the VOT of

Table 2. Mean VOT (*SD* in parentheses) by group, language, and segment. The values for /g/ are presented separately for devoiced and prevoiced productions.

		Dutch		German	
		Monolinguals	Bilinguals	Bilinguals	Monolinguals
	/k/	32 (21)	39 (22)	56 (30)	76 (28)
devoiced	/g/	26 (13)	31 (18)	23 (13)	19 (11)
prevoiced	/g/	−81 (57)	−88 (43)	−88 (40)	−61 (31)

these devoiced productions of /g/ largely overlapped with those of /k/. Table 2 provides the mean VOT and standard deviations (*SD*) by group, language, and plosive. The means and *SD* here and elsewhere in the Results section were obtained by first calculating values for each child and then averaging them to obtain the group mean and *SD*.

The Dutch monolingual children produced the most prevoiced productions of /g/ (27%, *SD*=28), followed by the bilingual children who prevoiced 13% (*SD*=15) of their /g/ tokens in Dutch and 13% (*SD*=21) in German. German monolingual children only prevoiced 3% (*SD*=6) of their /g/ tokens. Inspection of individual children's production reveals different patterns of prevoicing across and within groups. Only one Dutch monolingual child consistently prevoiced all /g/ tokens. The number of children who sometimes prevoiced /g/ was highest in the group of Dutch monolinguals and lowest in the group of German monolinguals. The bilinguals fell in between the two monolingual groups, with more prevoicing of /g/ tokens in Dutch than in German. The number of children who never prevoiced /g/ was lowest in the group of Dutch monolinguals, and highest in the group of German monolinguals. The bilingual children again fell between the two monolingual groups, with fewer bilingual children who never prevoiced /g/ speaking Dutch than speaking German. These prevoicing patterns are displayed in Table 3.

As expected based on previous research, Dutch monolingual and bilingual children in both languages produced prevoicing less frequently for /g/ than for /b/ and /d/ (Figure 2) because the production of prevoicing is increasingly difficult for places of articulation further back in the mouth (Cho & Ladefoged, 1999; Ohala & Riordan, 1979).

Analyses

Mixed-effects models were conducted in R software (R Core Team, 2013) using the lme4 package (Bates et al., 2015). Models were constructed using a maximal random effects structure (Barr et al., 2013). Specifically, random by-Child and by-Item intercepts model the variance between children or items (respectively) in prevoicing likelihood or VOT duration (depending on the model as outlined below). Based on previous research, which found no effect of age or language exposure on the production of 'voiced' plosives, these factors were not included in the present analyses (Stoehr et al., 2018). The *p*-values for *t*-statistics were obtained using the lmerTest package (Kuznetsova et al., 2017).

Table 3. Prevoicing patterns across groups: percentages and absolute numbers of children.

	Dutch		German	
	Monolinguals	Bilinguals	Bilinguals	Monolinguals
Always prevoiced	3% (1/30)	0% (0/28)	0% (0/28)	0% (0/29)
Sometimes prevoiced	77% (23/30)	71% (20/28)	57% (16/28)	31% (9/29)
Never prevoiced	20% (6/30)	29% (8/28)	43% (12/28)	69% (20/29)

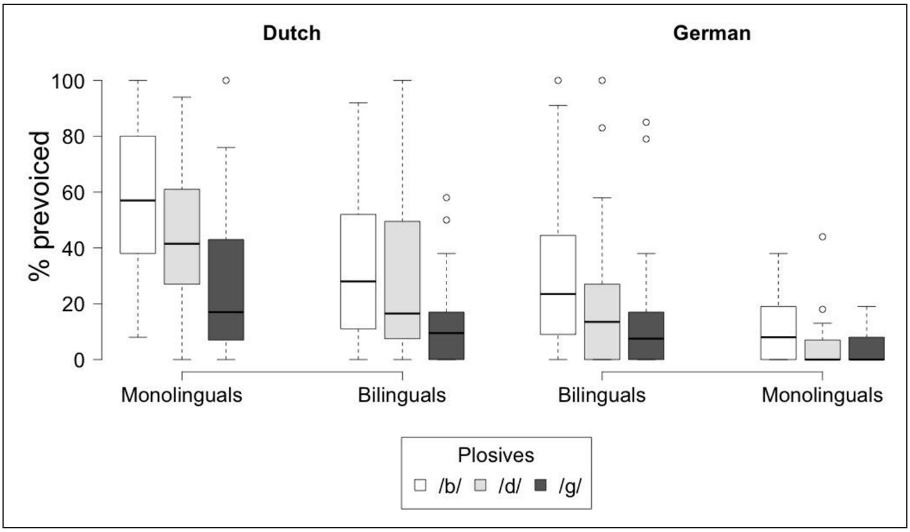

Figure 2. Percent Prevoiced /g/ by Language and Group Aggregated by Child. Results for /b/ and /d/ Are Based on Stoehr et al. (2018).

Dutch production of monolingual and bilingual children

> *RQ1a*. Do Dutch monolingual children and bilingual children alike generalize the feature [voice] to the dorsal place of articulation, resulting in instances of prevoiced /g/ in Dutch?

The likelihood of /g/ prevoicing in Dutch was compared between bilingual (13%) and monolingual children (27%) using a mixed-effects logistic regression model (612 observations). The dependent variable was Prevoicing: productions with a negative VOT were counted as prevoiced (coded as 1) and those with a positive VOT as devoiced (coded as 0). The difference between monolinguals and bilinguals was tested with a between-subjects independent variable, Language Background (monolingual = −1, bilingual = 1). In addition to the random intercepts for Child and Item, the model included by-Item random slopes for Language Background to allow the effect of language background to vary for each item, model R code: Prevoicing ~ Language Background + (1 | Child) +

(1 + Language Background | Item). The results confirmed that bilingual children prevoiced fewer /g/ tokens in Dutch than their Dutch monolingual peers ($\beta_{Language\ Background}=-0.617$, $SE=0.221$, $z=-2.789$, $p=.005$).

> *RQ1b*. When producing /g/ without prevoicing, do Dutch monolingual and bilingual children produce these devoiced instances of /g/ in line with their Dutch segment /k/?

The differences in VOT duration between devoiced /g/ and /k/ tokens produced in Dutch by bilingual and monolingual children were compared using a mixed-effects linear regression model (1234 observations). The continuous dependent variable was VOT in ms. VOT differences between devoiced /g/ and /k/ were tested with a within-subjects independent variable Voicing (/g/ = −1, /k/ = 1; each referring to the target pronunciation of a word) and the differences between monolinguals and bilinguals were tested with a between-subjects variable, Language Background (monolingual = −1, bilingual = 1). An interaction term addressed whether bilinguals and monolinguals differed in distinguishing /g/ and /k/ in production. In addition to the random intercepts for Child and Item, the model included by-Child random slopes for Voicing to allow the effect of voicing to vary for each child, and by-Item random slopes for Language Background to allow the effect of language background to vary for each item, model R code: VOT ~ Voicing × Language Background + (1 + Voicing | Child) + (1 + Language Background | Item). The model detected no difference in VOT duration between devoiced /g/ in a Dutch context and Dutch /k/ ($\beta_{Voicing}=-3.186$, $SE=2.887$, $t=-1.104$, $p>.250$) and no difference between monolinguals and bilinguals ($\beta_{Language\ Background}=2.938$, $SE=1.535$, $t=1.914$, $p=.063$; $\beta_{Language\ Background*Voicing}=-0.383$, $SE=1.117$, $t=-0.343$, $p>.250$).

German production of bilingual children and monolingual controls

> *RQ2a*. Do bilingual children prevoice /g/ more frequently in German than German monolinguals, providing evidence for feature-level cross-linguistic influence from Dutch, which lacks /g/ or do they produce /g/ like monolinguals?

The likelihood of /g/ prevoicing in German was compared between bilingual (13%) and monolingual children (3%) using a mixed-effects logistic regression model (802 observations). The model was identical to the one reported for RQ1a, model R code: Prevoicing ~ Language Background + (1 | Child) + (1 + Language Background | Item). The model confirmed that bilingual children prevoiced more /g/ tokens in German than monolinguals ($\beta_{Language\ Background}=0.809$, $SE=0.323$, $z=2.505$, $p=.012$).

> *RQ2b*. When producing /g/ in German without prevoicing, do bilinguals produce shorter VOT than monolinguals, providing evidence for feature-level cross-linguistic influence from Dutch, which lacks /g/ or do they produce /g/ like monolinguals?

VOT duration differences in devoiced /g/ by German bilinguals and German monolinguals were compared using a mixed-effects linear regression model (734 observations).

The continuous dependent variable was VOT in ms, and the remaining model structure was identical to the one reported in RQ2a, model R code: VOT ~ Language Background + (1|Child) + (1 + Language Background|Item). The model found that bilingual children produced devoiced instances of /g/ in German with longer VOT than monolinguals ($\beta_{Language\ Background}$ = 1.987, SE = 0.880, t = 2.257, p = .032).

Bilingual children in Dutch and German

> *RQ3a*. When producing /g/ with prevoicing, do bilingual children employ their German /g/ segment in Dutch or do they generalize their Dutch voicing feature to /g/ in production?

The likelihood of bilinguals prevoicing /g/ was compared in Dutch (13%) and German (13%), using a mixed-effects logistic regression model (708 observations). The dependent variable was Prevoicing as in the models on RQ1a and RQ2a. The difference between German and Dutch was tested on a within-subjects independent variable, Language (German = −1, Dutch = 1). In addition to the random intercepts for Child and Item, the model included by-Child random slopes for Language to allow the effect of language to vary for each child, model R code: Prevoicing ~ Language + (1 + Language|Child) + (1 |Item). The model did not detect any difference between Dutch and German in terms of the likelihood of prevoiced /g/ tokens ($\beta_{Language}$ = 0.386, SE = 0.282, z = 1.370, p = .171).

> *RQ3b*. When producing /g/ without prevoicing, do bilingual children employ their German /g/ segment in Dutch or do they generalize their Dutch voicing feature to /g/ in production?

The differences in the duration of bilinguals' VOT on devoiced /g/ (87% of all productions) in Dutch and German were compared using a mixed-effects linear regression model (612 observations). The continuous dependent variable was VOT in ms, and the remaining model structure was identical to the one reported in RQ3a, model R code: VOT ~ Language + (1 + Language|Child) + (1|Item). This analysis revealed that bilingual children produced longer VOT in devoiced /g/ in a Dutch compared with a German context ($\beta_{Language}$ = 4.278, SE = 1.673, t = 2.558, p = .021).

Discussion

This study investigated whether the production of /g/ by monolingual Dutch-speaking (hereafter, Dutch monolinguals) and Dutch–German speaking bilingual children (hereafter, bilinguals) reflects feature-level influence or segmental adaptation. We addressed this overarching question through three specific questions: first, do Dutch monolinguals generalize the feature [voice] to the dorsal place of articulation in Dutch, resulting in productions of prevoiced /g/, or do they produce non-Dutch /g/ in line with their Dutch segment /k/? Second, do bilinguals follow the same strategies as monolinguals in Dutch or do they transfer their German /g/ segment into Dutch? Third, is the production of /g/

in German by bilingual preschoolers influenced by their Dutch phonological system despite the lack of /g/ in Dutch or do they produce German /g/ like monolinguals? Our investigation of all three questions supports an interplay between feature generalization and segmental adaptation in children's language processing, as discussed below.

Monolingual and bilingual children's production of /g/ in Dutch: evidence for feature generalization and segmental adaptation

Production of /g/ in a Dutch context can reveal whether language processing by Dutch-speaking children is mediated by phonological features. Dutch monolinguals prevoiced approximately one-quarter of their /g/ tokens; this was less than the proportion of prevoiced /b/ and /d/ tokens reported in Stoehr et al. (2018). This production pattern is in line with the decrease in prevoicing production as the plosive place of articulation moves back in the mouth, which has been observed cross-linguistically in adults (Cho & Ladefoged, 1999; Ohala & Riordan, 1979).

Dutch monolingual children's production of prevoiced /g/ suggests generalization of the feature [voice], phonetically implemented with prevoicing in Dutch, to the dorsal place of articulation (Table 1, pattern A). Alternatively, one might argue that monolingual Dutch children resorted to production of /b/ or /d/ in place of /g/, which could be explained by segmental adaptation. However, as children were unambiguously producing a dorsal and not a labial or coronal segment, only the feature [voice] appears to have influenced children's production of /g/, making feature-level transfer a more parsimonious and precise account. The observed feature-level generalization in children's production extends parallel findings for production by adults (Nielsen, 2011; Olson, 2019) as well as perception by infants (e.g. Maye et al., 2008) and adults (Chládková et al., 2017; Finley & Badecker, 2009; Hestvik & Durvasula, 2016; Kraljic & Samuel, 2006; Lahiri & Reetz, 2002, 2010). Moreover, previous observations reporting /g/ in loanwords produced by Dutch-speaking adults (Hamann & de Jonge, 2015; Van Bezooijen & Gerritsen, 1994) could not exclude L2 transfer as the mechanism behind prevoicing in /g/. Here, the young age of the children excludes this explanation, further strengthening the interpretation that the feature [voice] was generalized within their Dutch phonological system.

The bilingual children prevoiced fewer /g/ tokens (13%) in a Dutch context than their Dutch monolingual peers (27%). This finding parallels the prevoicing patterns for Dutch /b/ and /d/, which bilingual children prevoice less frequently than Dutch monolinguals (Stoehr et al., 2018; see Kehoe et al., 2004 and Khattab, 2000 for German-Spanish and English-Arabic bilinguals). There are two possible explanations for the bilingual children's production of prevoiced /g/ in a Dutch context. First, they have been influenced by the same feature generalization process that appears to mediate Dutch monolingual children's production of prevoiced /g/ (Table 1, pattern A). Second, they may have borrowed /g/ from German, in which the bilingual children also prevoiced 13% of all /g/ tokens (Table 1, pattern C). If the bilinguals' production of /g/ in Dutch was the result of borrowing from German, their devoiced instances of /g/ should also have been produced in a more German-like manner, exhibiting longer VOT than Dutch monolinguals' production.

However, the predominantly devoiced /g/-initial items in Dutch of monolinguals and bilinguals (73% and 87% of all tokens, respectively) did not differ in terms of VOT duration. This counts against the argument that the bilinguals borrowed /ɡ/ from German. Yet, VOT durations of devoiced /ɡ/ produced by both groups were indistinguishable from the VOT durations of the ‘voiceless’ dorsal plosive /k/ in Dutch (Table 1, pattern B). By contrast, when the same children failed to prevoice Dutch /b/ and /d/, these devoiced productions had significantly shorter VOT than /p/ and /t/ (Stoehr et al., 2018). The fact that they maintained a voicing contrast at the labial and coronal but not at the dorsal place of articulation strengthens the argument that both monolingual and bilingual Dutch-speaking children adapted non-Dutch /ɡ/ to their already-existing Dutch /k/ category (Silverman, 1992). Monolingual and bilingual children’s production of prevoiced versus devoiced /ɡ/ in Dutch may thus result from two different mechanisms.

Bilingual children, unexpectedly, produced devoiced /ɡ/ with longer VOT in Dutch than in German. This difference provides further evidence against the hypothesis that bilingual children borrowed the /ɡ/ segment from German when producing (devoiced) /ɡ/ in Dutch. More specifically, if the bilingual children had aimed to produce their German /ɡ/ in a Dutch context, but failed to do so with prevoicing for articulatory reasons (Cho & Ladefoged, 1999; Ohala & Riordan, 1979), they should have produced similar VOT durations in Dutch and in German. These results, conversely, further strengthen our previous conclusions regarding how monolingual as well as bilingual children produce /ɡ/ in Dutch: while their prevoiced /ɡ/ is best explained by feature generalization, their devoiced /ɡ/ with VOT indistinguishable from Dutch /k/ appears to result from adapting /ɡ/ in a Dutch context to /k/.

In sum, productions of /ɡ/ in a Dutch context by both monolingual and bilingual children provide evidence for the existence of features in their still-developing phonology. Both groups of children followed the same patterns, either combining two native features [voice] and [dorsal] into a novel segment /ɡ/ or adapting /ɡ/ to the ‘voiceless’ segment /k/, which shares the [dorsal] place feature. Most children applied both strategies, likely because uncertainty, caused by little experience with /ɡ/ did not allow for automatization. Uncertainty in the categorization of non-native segments is also common in adult speech perception, where non-native segments are usually considered to be categorized, that is, mapped onto an L1 segment, when response consistency exceeds 70% (e.g. Tyler et al., 2014). This, in turn, implies that inconsistent responses in 30% of cases should be expected when (adult) listeners deal with non-native segments.

It must be acknowledged that there were relatively few prevoiced productions in Dutch by monolinguals (27%) and bilinguals (13%). One might argue that this pattern is due to typical variation rather than the result of feature generalization. However, if prevoiced /ɡ/ was the result of typical variation, it should have been found in German monolingual children as well, which was not the case. Even though the prevalence of prevoiced productions was low, their mere presence appears to be driven by feature generalization within the children’s Dutch phonological system.

Bilingual children’s production of /ɡ/ in German: tentative evidence for cross-linguistic influence at the feature-level

Comparing bilingual and German monolingual children’s production of German /ɡ/ can provide insight into the nature of cross-linguistic influence. Previous research has reported cross-linguistic influence 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, 2002; Kehoe et al., 2004; Khattab, 2000). These studies, however, only included production of segments that were present in both of the children's languages and therefore could not address whether cross-linguistic influence operates on features.

The present study showed that bilingual children prevoiced more /g/ tokens than German monolinguals, who virtually never prevoiced /g/ (Table 1, pattern D). This consistently higher proportion of prevoiced /g/ tokens in the bilingual children's heritage language German cannot be explained by cross-linguistic influence 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 by prevoicing, seems to have influenced bilingual children's production of German /g/ in the same way it influenced their German /b/ and /d/ (Stoehr et al., 2018).

In the present data, cross-linguistic influence of Dutch on German caused the bilingual children to produce articulatorily *more* complex prevoicing, which is typically only acquired later by monolingual children (Kewley-Port & Preston, 1974; Macken & Barton, 1980) and is particularly difficult to produce at the dorsal place of articulation (Cho & Ladefoged, 1999; Ohala & Riordan, 1979). These findings corroborate previous findings on adult L2-English speakers whose native Arabic dialect lacks /p/, and who produced English /ptk/ without aspiration, based on their L1 phonology (Flege & Port, 1981). 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 cross-linguistic influence at the feature-level.

When bilingual children produced /g/ in German without prevoicing, they did so with longer VOT than the German monolinguals. This is unexpected and cannot be explained by either feature-level cross-linguistic influence or segmental adaptation from Dutch. To understand why bilinguals produced devoiced /g/ in German with longer VOT than monolinguals, we reanalyzed the same children's production of devoiced /b/ and /d/ in German as reported in Stoehr et al. (2018). It turned out that bilinguals also produced devoiced /b/ and /d/ with longer VOT than monolinguals, showing that their productions of devoiced /g/ are in fact in line with their German phonological system. In sum, when cross-linguistic influence occurs, it leads to production of prevoiced /g/ in the bilinguals' German, which is indicative of feature-level cross-linguistic influence. Those productions unaffected by cross-linguistic influence fall in line with the typical production pattern of this bilingual population.

Establishing language proficiency and directions for future research

The bilingual children's language proficiency was determined by parental report rather than measured using linguistic tasks. The parents estimated the children to be more proficient in Dutch (the majority language) than in German (the heritage language). If parents overestimated their bilingual child's proficiency in German, the bilingual and monolingual Dutch children's linguistic backgrounds may have been more similar than assumed. All bilingual children, however, were able to complete both naming tasks in German, produced language-specific VOT for devoiced /g/ tokens, and produced fewer prevoiced /g/ tokens in Dutch than monolingual Dutch children. These observations and results confirm the children's bilingual status and demonstrate that the bilingual children were not identical to their Dutch monolingual peers. Nevertheless, the bilingual children were more proficient in Dutch than in German. Cross-linguistic influence of Dutch prevoicing on German /g/ might only occur for children who are more proficient in Dutch – the prevoicing language. Along these same lines, one could speculate that segment-level influence of German /g/ on Dutch could occur for bilingual children who are more proficient in German – the language with a native /g/ segment. These hypothetical scenarios warrant further investigation of the interplay between phonological representations, language proficiency, and cross-linguistic influence. However, they do not negate this study's evidence for feature-level cross-linguistic influence in bilingual children.

Conclusion

This study provides evidence for the existence of phonological features within preliterate children's native phonology. Three main findings demonstrate that features play a crucial role in children's speech production:

1. Dutch monolingual children generalized the feature [voice] to the dorsal place of articulation, suggesting that features play a role in children's language processing. This generalization, however, was not consistent as Dutch-speaking children more often resorted to producing the perceptually and featurally similar native segment /k/ as a result of segmental adaptation. Inconsistent use of two strategies may result from lack of automatization in the production of non-native segments.
2. Bilingual children appeared to be following the same production patterns for /g/ in their majority language Dutch as their monolingual Dutch-speaking peers, although they were familiar with /g/ from their heritage language German.
3. Bilingual children's production of /g/ in their heritage language German reflects phonological influence by the majority language Dutch, despite the absence of /g/ in the Dutch phoneme inventory.

These results are an important contribution to the field of speech processing as they demonstrate that speech production is mediated by features. This finding was obtained by testing monolingual and bilingual children speaking Dutch or Dutch and German, and therefore complements previous results from monolingual speakers of English. These results have crucial implications for our understanding of bilingualism, demonstrating that bilingual children generalize features within the language system they are currently using instead of resorting to their other language. The bilingual children's production of /g/ further suggests that the phonological systems of bilingual children may be more prone to cross-linguistic influence from the majority to the heritage language than vice versa. Importantly, if such cross-linguistic influence occurs, it occurs between features, and not just on segments shared by a bilingual's two languages. These results provide important theoretical insight into the nature of cross-linguistic influence.

Acknowledgements

We thank Rianne van den Berghe, Arjan Cuppen, Lottie Gort, Eva Koch, and Jana Loh for their help collecting the data; Franziska Klier, Ann-Katrin Ohlerth, and Natascha Roos for their help annotating the data; and Magda Altman for proofreading this manuscript. We are also grateful to two anonymous reviewers for insightful comments and suggestions on an earlier version of this manuscript.

CrediT statement

Antje Stoehr: Conceptualization, Methodology, Formal analysis, Data curation, Investigation, Writing – Original Draft, Visualization, Project administration, Funding acquisition. **Titia Benders**: Conceptualization, Methodology, Formal analysis, Writing – Review & Editing, Supervision. **Janet van Hell**: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Funding acquisition. **Paula Fikkert**: Conceptualization, Methodology, Writing – Review & Editing, Supervision.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This work was supported by the Basque Government [BERC 2018-2021 program]; the Spanish State Research Agency [BCBL Severo Ochoa excellence accreditation SEV-2015-0490]; the European Union's Horizon 2020 research and innovation program [Marie Skłodowska-Curie grant 843533]; and the National Science Foundation [BCS1349110; OISE 1545900].

ORCID iD

Antje Stoehr https://orcid.org/0000-0002-6065-5952

Supplemental material

Supplemental material for this article is available online.

Notes

1. Hereafter, the terms 'voiced' and 'voiceless' in single quotation marks are used to refer to phonologically voiced and voiceless plosives, respectively.
2. There are different views on the featural representation of voicing contrasts. While some researchers assume that the voicing contrast is represented by a single feature [voice] (e.g. Lahiri, 2018; Lombardi, 1995), others argue that it is best described by monovalent features that differ depending on phonetic implementations (e.g. Iverson & Salmons, 1995). In the latter view, prevoicing languages like Dutch have the feature [voice], while aspiration languages like German have the feature [spread glottis]. For the sake of simplicity, we use the notation [voice] to refer to both Dutch and German voiced plosives in this article.
3. The Dutch and German items were unbalanced, with six familiar words in German versus one English loanword and three names in Dutch. These names were introduced to the children by the experimenter, who produced the names with prevoicing, while the existing words were never produced by the experimenter. Should children have imitated the experimenter's

production, Dutch /g/ tokens should have been prevoiced more frequently than /b/ and /d/ tokens, but the opposite was the case. Likewise, bilinguals should have prevoiced /g/ more frequently in Dutch than in German, which was not the case in the present data. The unbalanced stimuli and the presentation format of the names in Dutch thus seem to not have unduly affected the results.

4. Included and excluded data points by group and plosive: bilinguals in Dutch: /g/: 320 included, 41 excluded (11.4%), /k/: 370 included, 19 excluded (4.9%); Bilinguals in German: /g/: 388 included, 34 excluded (8.1%), /k/: 412 included, 54 excluded (11.6%); Dutch monolinguals: /g/: 292 included, 88 excluded (23.2%), /k/: 375 included, 78 excluded (17.2%); German monolinguals: /g/: 414 included, 29 excluded (6.6%), /k/: 391 included, 42 excluded (9.7%).

References

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

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://10.18637/jss.v067.i01

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

Braunschweiler, N. (1997). Integrated cues of voicing and vowel length in German: A production study. *Language and Speech*, *40*(4), 353–376. https://doi.org/10.1177/002383099704000403

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). Benjamins.

Chládková, K., Podlipský, V. J., & Chionidou, A. (2017). Perceptual adaptation of vowels generalizes across the phonology and does not require local context. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(2), 414–427. https://doi.org/10.1037/xhp0000333

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, *27*(2), 207–229. https://doi.org/10.1006/jpho.1999.0094

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

Clements, G. N. (1985). The geometry of phonological features. *Phonology*, *2*(1), 225–252. https://doi.org/10.1017/S0952675700000440

Clements, G. N. (2003). Feature economy in sound systems. *Phonology*, *20*(3), 287–333. https://doi.org/10.1017/S095267570400003X

Clements, G. N. (2005). The role of features in speech sound inventories. In E. Raimy & C. Cairns (Eds.), *Contemporary views on architecture and representations in phonological theory* (pp. 19–68). MIT Press.

Clements, G. N., & Hume, E. (1995). The internal organization of speech sounds. In J. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 245–306). Blackwell.

Cristià, A., Seidl, A., & Francis, A. L. (2011). Phonological features in infancy. 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). Benjamins.

Deuchar, M., & Clark, A. (1996). Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics*, *24*(3), 351–365. https://doi.org/10.1006/jpho.1996.0019

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

Durand, J., & Laks, B. (2002). *Phonetics, phonology and cognition*. Oxford University Press.

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*(2), 148–163. https://doi.org/10.3109/02699206.2011.595526

Finley, S., & Badecker, W. (2009). Artificial language learning and feature-based generalization. *Journal of Memory and Language*, *61*(3), 423–437. https://doi.org/10.1016/j.jml.2009.05.002

Flege, J. E., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech*, *24*(2), 125–146. https://doi.org/10.1177/002383098102400202

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*(2), 1025–1032. https://doi.org/10.1121/1.1536169

Grimm, H., & Doil, H. (2000). *ELFRA: Elternfragebögen für die Früherkennung von Risikokindern, Handanweisung* [Parental questionnaire on early diagnosis of at-risk children]. Hogrefe.

Gussenhoven, C., & Jacobs, H. (2017). *Understanding phonology*. Routledge.

Halle, M. (1964). On the bases of phonology. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language: Readings in the philosophy of language* (pp. 324–333). Prentice-Hall.

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*. University of Glasgow. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0950.pdf

Hestvik, A., & Durvasula, K. (2016). Neurobiological evidence for voicing underspecification in English. *Brain and Language*, *152*, 28–43. https://doi.org/10.1016/j.bandl.2015.10.007

Iverson, G. K., & Salmons, J. C. (1995). Aspiration and laryngeal representation in Germanic. *Phonology*, *12*(3), 369–396. http://www.jstor.org/stable/4420084

Iverson, G. K., & Salmons, J. C. (2005). Filling the gap: English tense vowel plus final /š/. *Journal of English Linguistics*, *33*(3), 207–221. https://doi.org/10.1177/0075424205282561

Jakobson, R. (1941). *Kindersprache, Aphasie und allgemeine Lautgesetze* [Child language, aphasia, and phonological universals]. Almqvist & Wiksell.

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

Johnson, C., & Wilson, I. (2002). Phonetic evidence for early language differentiation: Research issues and some preliminary data. *International Journal of Bilingualism*, *6*(3), 271–289. https://doi.org/10.1177/13670069020060030401

Kager, R., van der Feest, S., Fikkert, P., Kerkhoff, A., & Zamuner, T. S. (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. 41–90). Benjamins.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, *60*(2), 286–319. https://doi.org/10.2307/413642

Kehoe, M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German–Spanish children. *Bilingualism: Language and Cognition*, *7*(1), 71–88. https://doi.org/10.1017/S1366728904001282

Kewley-Port, D., & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics*, *2*(3), 194–210. https://doi.org/10.1016/S0095-4470(19)31270-7

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* (Vol. 8, pp. 95–122). http://www.leeds.ac.uk/arts/homepage/387/leeds_working_papers_in_linguistics_and_phonetics

Kiparsky, P. (1995). The phonological basis of sound change. In J. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 640–670). Blackwell.

Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review*, *13*, 262–268. https://doi.org/10.3758/BF03193841

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lahiri, A. (2018). Predicting universal phonological features. In L. Hyman & F. Plank (Eds.), *Phonological typology* (pp. 229–272). Mouton de Gruyter.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. *Laboratory Phonology*, *7*, 637–676. https://doi.org/10.1515/9783110197105.2.637

Lahiri, A., & Reetz, H. (2010). Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics*, *38*(1), 44–59. https://doi.org/10.1016/j.wocn.2010.01.002

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word*, *20*(3), 384–422. https://doi.org/10.1080/00437956.1964.11659830

Lombardi, L. (1995). Laryngeal neutralization and syllable wellformedness. *Natural Language and Linguistic Theory*, *13*(1), 39–74. http://www.jstor.org/stable/4047887

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*(3), 433–458. https://doi.org/10.1017/S0305000900002774

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252–272. https://doi.org/10.1016/j.jml.2007.03.005

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, *11*(1), 122–134. https://doi.org/10.1111/j.1467-7687.2007.00653.x

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). Benjamins.

Mielke, J. (2008). *The emergence of distinctive features* [Unpublished doctoral dissertation]. The Ohio State University.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, *39*(2), 132–142. https://doi.org/10.1016/j.wocn.2010.12.007

Ohala, J. J. (1981). The listener as a source of sound change. In C. Masek, R. Hendrik, & M. Miller (Eds.), *Papers from the Parasession on language and behavior* (pp. 178–203). Chicago Linguistic Society.

Ohala, J. J. (2010). The relation between phonetics and phonology. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 653–677). John Wiley.

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). Acoustical Society of America.

Olson, D. J. (2019). Feature acquisition in second language phonetic development: Evidence from phonetic training. *Language Learning*, *69*(2), 366–404. https://doi.org/10.1111/lang.12336

Pater, J., Stager, C., & Werker, J. (2004). The perceptual acquisition of phonological contrasts. *Language*, *80*(3), 384–402. https://doi.org/10.1353/lan.2004.0141

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

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). Lawrence Erlbaum.

Silverman, D. (1992). Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology*, *9*(2), 289–328. http://www.jstor.org/stable/4420058

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*(1), 94–109. https://doi.org/10.1111/desc.12097

Stager, C., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381–382. https://doi.org/10.1038/41102

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*(4), 483–518. https://doi.org/10.1177/0267658317704261

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2018). Heritage language exposure impacts voice onset time of Dutch-German simultaneous bilingual preschoolers. *Bilingualism: Language and Cognition*, *21*(3), 598–617. https://doi.org/10.1017/S1366728917000116

Stoel-Gammon, C., Williams, K., & Buder, E. (1994). Cross-language differences in phonological acquisition: Swedish and American /t/. *Phonetica*, *51*(1–3), 146–158. https://doi.org/10.1159/000261966

Swingley, D., & Aslin, R. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*(2), 147–166. https://doi.org/10.1016/s0010-0277(00)00081-0

Szagun, G., Stumper, B., & Schramm, S. A. (2009). *FRAKIS – Fragebogen zur frühkindlichen Sprachentwicklung. PSYNDEX Tests Info* [Questionnaire on early language development]. Pearson.

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, *71*(1), 4–21. https://doi.org/10.1159/000356237

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition*, *16*(1), 86–110. https://doi.org/10.1017/S1366728912000284

Van Bezooijen, R., & Gerritsen, M. (1994). De uitspraak van uitheemse woorden in het Standaard-Nederlands: een verkennende studie [The pronunciation of foreign words in Standard Dutch: An exploratory study]. *De Nieuwe Taalgids*, *87*, 145–160. https://www.dbnl.org/tekst/_taa008199401_01/_taa008199401_01_0014.php

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Age and vocabulary size influences on the phonological representation of newly learned words in infants aged 14 to 20 months. *Infancy*, *3*(1), 1–30. https://doi.org/10.1207/S15327078IN0301_1

Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Science*, *30*(4), 553–566. https://doi.org/10.1037/0012-1649.30.4.553

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). Multilingual Matters.

Zink, I., & Lejaegere, M. (2002). *N-CDIs: Lijsten voor Communicatieve Ontwikkeling. Aanpassing en hernormering van de MacArthur CDIs van Fenson et al. 1993* [Lists for communicative development. Adaptation and standardization of the MacArthur CDIs of Fenson et al. 1993]. Acco.