

Linking Variation in Perception and Production in Sound Change: Evidence from Dutch Obstruent Devoicing

Language and Speech
2020, Vol. 63(3) 660–685
© The Author(s) 2019



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0023830919880206
journals.sagepub.com/home/las



Anne-France Pinget 

René Kager

Utrecht Institute of Linguistics OTS, Utrecht University, The Netherlands

Hans Van de Velde 

Utrecht Institute of Linguistics OTS, Utrecht University, The Netherlands;
Fryske Akademy, Leeuwarden, The Netherlands

Abstract

This study investigates the link between the perception and production in sound change in progress, both at the regional and the individual level. Two devoicing processes showing regional variation in Dutch are studied: the devoicing of initial labiodental fricatives and of initial bilabial stops. Five regions were selected, to represent different stages of change in progress. For each region, 20 participants took part in production (Study 1) and perception (Study 2) experiments. First, the results of the production tasks give additional insight in the regional and individual patterns of sound change. Second, the regional perceptual patterns in fricatives match the differences in production: perception is the most categorical in regions where the devoicing process is starting, and the least categorical in regions where the process of devoicing is almost completed. Finally, a clear link is observed between the production and perception systems undergoing sound change at the individual level. Changes in the perceptual system seem to precede changes in production. However, at the sound change completion, perception lags behind: individuals still perceive a contrast they no longer produce.

Keywords

Speech production, speech perception, sound change, devoicing processes, obstruents

Corresponding author:

Anne-France Pinget, Utrecht Institute of Linguistics OTS, Utrecht University, Trans 10, Utrecht, 3512 JK, The Netherlands.

Email: a.c.h.Pinget@uu.nl

Introduction

There are different views about the link that exists between the processes of perception and production in the human cognitive system. This paper investigates the link between the speech perception and production systems in the context of sound change in progress, and presents evidence for the view that perception and production are intrinsically linked, and that any changes in production are accompanied by changes in perception, and the other way around. The case of sound change is particularly informative, since the individual production and perception systems are together involved in a process with a start and an end point. The production and perception studies of the current paper are conducted on the same pool of participants stratified for region, reflecting different stages of change in progress, allowing the investigation of the link between production and perception at both the group and individual levels.

Motor Theory (Lieberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman, Harris, Hoffman, & Griffith, 1957; Liberman & Mattingly, 1985; Liberman & Whalen, 2000) and Direct-Perception Theory (Fowler, 1986; Goldstein & Fowler, 2003) are examples of theories positing a clear link between speech production and perception. Motor Theory argued that adults perceive speech by making reference to their own articulation. They decode the speech signal by recovering the movements that have produced it. In the same way, Direct-Perception Theory suggested that listeners directly perceive the gestures or productions of the speaker. Within the framework of exemplar-based theories (e.g., Goldinger, 1997, 1998; Hawkins, 2003; Johnson, 1997; Pierrehumbert, 2001, 2002, 2003; Pisoni, 1997) some scholars supported the idea that representations for speech perception and speech production might be intrinsically different. Speech perception is highly malleable and this great malleability is shown to be essential for comprehension (e.g., Sumner & Samuel, 2009). Production, on the contrary, is less flexible. Listeners may perfectly understand and process variants they cannot produce themselves (Thomas, 2000). Garrett and Johnson (2013) proposed a dual representation model in which perception and production are based on different sets of exemplars. Since the space of familiar exemplars for speech perception is necessarily larger and more diverse than the space of exemplars used in speech production, they distinguished between the two, proposing one phonetic space for listening and another one for speaking. In the current study, we assume, as in the abovementioned frameworks, that the perception of phonetic contrasts necessarily includes some degree of categoricalness, even if the contrasts are inherently gradient.

To date, quite a number of studies have investigated the link between perception and production. These studies differ substantially in their methodological approach. Some studies investigated whether adults exhibit a perceptual preference for phonetic tokens comparable to those they typically produce (e.g., Harrington, Kleber, & Reubold, 2012; Kleber, Harrington, & Reubold, 2012). Others look at the relationship between listeners' judgments on category goodness and their own production (e.g., Evans & Iverson, 2004; Frieda, Walley, Flege, & Sloane, 2000; Johnson, Flemming, & Wright, 1993; Nakai, 1998; Newman, 1997, 2003). An alternative type of evidence comes from studies showing that different groups of speakers employ different articulatory strategies in achieving the same phonological goals (e.g., Bell-Berti, Raphael, Pisoni, & Sawusch, 1979; Ferguson & Quené, 2014; Fox, 1982).

The results of these studies often seem to be contradictory. Some studies found evidence in favor of a clear link between production and perception (e.g., Bell-Berti *et al.*, 1979; Evans & Iverson, 2007; Fox, 1982; Fridland & Kendall, 2012; Harrington, Kleber, & Reubold, 2008; Janson, 1983; Kendall & Fridland, 2012; Kleber *et al.*, 2012; Perkell *et al.*, 2004), whereas other studies showed evidence for a link exclusively at a very abstract level (Mitterer & Ernestus, 2008) or even no link at all between perception and production (Ainsworth & Paliwal, 1984; Bailey & Haggard, 1973;

Grosvald, 2009; Grosvald & Corina, 2012; Paliwal, Lindsay, & Ainsworth, 1983). Bell-Berti et al. (1979) demonstrated that differences in production strategies for the /i/ vowel were significantly correlated with differences in perception. Looking at coarticulation, Harrington et al. (2008) and Kleber et al. (2012) found that speakers who produced vowels with fewer coarticulatory influences also perceptually adjusted less for these influences. Perkell et al. (2004) found that speakers who articulated phonemically contrasting vowels more distinctly also showed a greater ability to distinguish vowel contrasts. Besides evidence for the hyperspace effect (see Johnson et al., 1993), Newman (2003) also found significant, though weak, correlations between listeners' perceptual prototypes and their average voice onset times (VOT) for English stop consonants. Listeners who selected as a perceptual prototype a stop consonant with a longer VOT were likely to show longer VOTs in their production of the same consonants. In contrast to this evidence for a direct link, Bailey and Haggard (1973) tested correlations between production and perception measures of English initial stops and found no correlation between average VOTs produced in voiced and voiceless consonants and listeners' perceptual category boundaries for these consonants. Similarly, Ainsworth and Paliwal (1984) and Frieda et al. (2000) found no correlation between performance on production and perception tasks for English glides and the English /i/ vowel respectively. Grosvald (2009) showed that participants who produce more extensive coarticulation do not perceive differences between coarticulated variants more accurately. Exploring to what degree a speaker's own productive system mirrors or interacts with his/her perceptual system appears to be complex.

This paper reports on two studies that tackle the question of the existence of a link between speech perception and production in the specific context of sound change in progress. Only a small number of studies examined the link between production and perception in sound change. In this type of research, most researchers used the apparent time construct: different age groups reflecting the stages of change. Beddor and Coetzee (2014) investigated whether speakers of Afrikaans who produce more innovative variants also weight this innovative property more heavily in perception. They tested two age groups of Afrikaans-speaking women and found a weak correlation ($r = .27$, $p = .01$) between production and perception of these variants. Fridland and Kendall (2012) and Kendall and Fridland (2012) looked at vowel changes in the frame of the American English Southern Vowel Shift (SVS) and Northern Cities Shift (NCS) by comparing regions instead of age groups. They acoustically analyzed participants' vowel productions and compared them to the same speakers' performance on the vowel identification task. Results showed that both regional affiliation and individual participation in regional shifts in production play a role in perception, indicating a clear relationship between the two processes. The current studies also tap into regional variation in order to look at perception and production during sound change.

In this paper we investigate two sound changes in consonantal contrasts. Until now, studies of this type have largely focused on vowels—often involved in chain shifts—and it is unclear whether the same kind of link can be found in consonantal contrasts. Furthermore, the perception of most consonants has been shown to be more categorical than the perception of vowels (e.g., Fry, Abramson, Eimas, & Liberman, 1962; Pisoni, 1975; Repp, 1984), which makes them particularly interesting in a study of the link between perception and production. We will seek an answer to the following question:

RQ 1: How do patterns of variation and change in speech production relate to speech perception?

In the case of evidence for a clear relationship between perception and production, it is important to investigate whether the two processes change simultaneously and which of the two processes takes the lead in the change. Ohala (1981) argued that misperception is an important source for the

origin of language change. Harrington et al. (2012) and Kleber et al. (2012) have also put forward the idea that in sound change, alternations in speech perception precede alternations in production. The reasoning is that the first step towards adopting a new form must be to perceive it. In this way, the language user first incorporates the new form in his/her perception and only later, if at all, the speaker will start to use the form in production. Harrington et al. (2012) and Kleber et al. (2012) even implied a causal relation between changes in perception and changes in production: “During a sound change in progress, the association between the perception and production of coarticulation passes through an unstable state during which the two modalities are out of alignment and in which changes to the coarticulatory relationships in perception lead those in production” (Kleber et al., 2012: 401). However, others argued for the opposite process: alternations of speech production precede alternations in perception. Janson (1983) found in a study of vowel change in Swedish that change in production precedes change in perception: “What seems to be clear is that, for an individual in a situation of change, perception seems to lag behind production. The old pattern of perception is still needed, even if one’s own production is changed” (Janson, 1983: 31). Some researchers also suggested that speech perception is not always affected in the case of a change in production (Evans & Iverson, 2007; Kraljic, Brennan, & Samuel, 2008). Listeners might also use phonetic cues in perception that they do not use in their own production, but that are used by other speakers in the community (Hay, Warren, & Drager, 2006; Thomas, 2000). Faced with these contradictory conclusions, we try to find additional evidence in order to answer the question:

RQ 2: Does change in speech perception precede or follow change in speech production?

These research questions are investigated in two studies: Study 1 focuses on speech production patterns (reported in section 3) and Study 2 focuses on speech perception (reported in section 4). In section 5, the insights of Study 1 and 2 are combined in order to obtain insight into the relationship between perception and production.

2 Methodology

To be able to answer our research questions, a language variety needs to be selected that displays changes in progress from the incipient phase to (nearly) full completion. Two variables from standard Dutch were chosen: the devoicing of the voiced labiodental fricative in word onset position (*v*) and the devoicing of the voiced bilabial stop in word onset position (*b*). The former one is a sound change in an advanced phase, showing regional variation in standard Dutch, and passing well above the level of consciousness in some parts of the language area, but not (yet) in others. The second one is an only recently observed variation pattern in the standard Dutch language that might be an incipient change (i.e., beginning phase of a change in progress). The language situation in the Dutch language area is explained in section 2.1 and the variables are described in section 2.2. Two cross-sectional studies were designed with a group level represented by five different regions and an individual level represented by 100 participants nested under these five regions. The regions represent different stages of the two sound changes in progress. The regions are described in section 2.3 and the participants in section 2.4.

2.1 Standard Dutch

Dutch is an official language in the Netherlands, Belgium and Surinam. In this paper we focus on the endogenous language area in the Low Countries, and leave Dutch as spoken in the former

Dutch colonies out of consideration. The variety of Dutch spoken in Flanders (the northern part of Belgium) is often labeled as Flemish, but in 1974 it was settled in a decree that Dutch is the name of the language spoken in Flanders. Already since the second half of the 19th century, the standardization policy towards the local varieties spoken in Flanders was one of integration with Dutch as spoken and written in the Netherlands. Meanwhile, Dutch has developed in a full-grown pluricentric language (Geeraerts & Van de Velde, 2013; Van de Velde et al., 2010). Flanders also moved from a diglossic to a diaglossic situation (Auer, 2005). Standardization has been very strong in the field of pronunciation, but it is at the phonetic level that strong divergence has been demonstrated among speakers of the standard variety (Van de Velde, van Hout, & Gerritsen, 1997; Van de Velde et al., 2010), but none of them leads to a different phoneme inventory. This is different from the local dialects, which are also phonologically different from standard Dutch (in addition to large lexical, morphological and syntactic differences; see Taeldeman & Hinskens, 2013). In the Netherlands, the uniform standard variety (of which the existence has been contested) has been replaced by a more variable colloquial standard. In Flanders, a colloquial variety has developed and has become the most frequently used variety in informal domains, at the cost of the local dialects. This colloquial variety is much further removed from the standard register than in the Netherlands, on the phonetic, lexical, morphological and syntactic level (Geeraerts & Van de Velde, 2013).

In this paper we focus on standard Dutch, the supra-regional variety that is spoken in formal situations (see section 3.1 for the data elicitation method). We follow Smakman's (2006: 283) inclusive interpretation of standard Dutch:

[T]he type of Dutch that avoids certain marked articulatory, lexical and grammatical structures. It is spoken in situations where people with various backgrounds come together and need to communicate effectively and impartially (shops, schools, in the professional world, and so on). This inclusive type of Standard Dutch contains both regional and non-regional traces but no dialect features, most importantly those that impair comprehension.

The differences between Flemish and Netherlandic Dutch are of the same order as those between American and British English: clear and increasing phonetic differences (speakers are immediately identified as being Dutch or Flemish (Van de Velde & Houtermans, 1999)) and lexical and grammatical differences that tend to become smaller (Geeraerts, Grondelaers, & Speelman, 1999). In both the Netherlands and Flanders, systematic regional differences show up in standard pronunciation, even among professional speakers such as high school teachers of the Dutch language (e.g., Kissine, Van de Velde, & Van Hout, 2005; Adank, Van de Velde, & Van Hout, 2007; Van de Velde et al., 2010; Van der Harst, Van de Velde, & Van Hout, 2014), and the regional background of speakers is reasonably well recognized by listeners (Pinget, Rotteveel, & Van de Velde, 2014).

2.2 The variables

2.2.1 Labiodental fricatives: (v) and (f). Standard Dutch is traditionally described as having a phonological distinction between voiced and voiceless fricatives (Booij, 1995). The major cue for the voiced/voiceless distinction is the presence or the absence of vocal cord vibration in the fricative (Slis & Cohen, 1969; van den Berg, 1988). Moreover, voiceless fricatives tend to be longer (Debrock, 1977; Slis & Cohen, 1969; Slis & van Heugten, 1989) and louder (Kissine, Van de Velde, & van Hout, 2003) than their voiced counterparts. During the last decades it has been frequently observed that word-initial voiced fricatives in standard Dutch are increasingly produced as voiceless because of a change in voicing, duration and/or intensity (Cassier & Van de Craen, 1986;

Cohen, Ebeling, Fokkema, & Holk, 1961; Donaldson, 1983; Goossens, 1974; Gussenhoven, 1999; Gussenhoven & Bremmer, 1983; Hamann & Sennema, 2005; Kissine et al., 2003, 2005; Mees & Collins, 1982; van der Wal & van Bree, 1992; Van de Velde, 1996; Van de Velde, Gerritsen, & van Hout, 1996). They also observed regional differences in the devoicing of voiced fricatives. Slis and van Heugten (1989) found stronger devoicing in the west of the Netherlands than in the south. Van de Velde et al. (1996) showed in a real-time study of radio broadcasters that fricative devoicing is a rapidly advancing change in progress in the Netherlands, and found the first signs of fricative devoicing in Flanders. In a follow-up study, these insights were refined by focusing on regional differences within the Netherlands and Flanders and on the implementation of the /v/-/f/ contrast (Kissine et al., 2003, 2005). They found that West-Flanders is the most conservative region, showing the highest scores for voicing of /v/, and that the north of the Netherlands is the most advanced with almost complete devoicing. Other regions exhibited intermediate states. In an electrolaryngographic study the devoicing of fricatives in Flanders was confirmed by Verhoeven and Haegeman (2007). Remarkably, the devoicing of voiced fricatives in Flanders is still unnoticed by Flemish language users, who consider devoicing of voiced fricatives as typically Netherlandic Dutch. In conclusion, there is large consensus and ample evidence that Dutch labiodental voiced fricatives in onset position are undergoing a sound change. It seems to be resulting in a merger or near-merger (Labov, 1994). This change shows important regional variation. Overall it can be considered as advanced, but not completed at this moment.

2.2.2 Bilabial stops: (b) and (p). Cross-linguistically, the main acoustic cue to distinguish between voiced and voiceless stops is voice onset time (VOT). Dutch differs from other Germanic languages in the way that the voicing distinction is phonetically implemented. Whereas some other Germanic languages contrast voiceless unaspirated and voiceless aspirated plosives, the contrast in Dutch is made between voiced and voiceless unaspirated plosives. In standard Dutch, voiced plosives are produced with a negative VOT and voiceless plosives are produced with little or no aspiration (Cohen et al., 1961; Keating, 1984; Rietveld & Van Heuven, 2009: 134–135). In a phonetic study of production and perception, Van Alphen and Smits (2004) investigated the voicing distinction in Dutch initial bilabial and alveolar plosives of 10 native speakers of Dutch from the Netherlands. They collected read speech material of monosyllabic Dutch words and showed that speakers generally implemented the contrast in terms of the presence or absence of prevoicing (negative VOT). Perceptually, prevoicing was also the strongest cue that listeners used to classify plosives as voiced or voiceless. Besides prevoicing, a range of secondary cues (e.g., the burst duration, the intensity of the burst) appeared to play some role in the contrast. Van Alphen and Smits (2004) noticed that prevoicing was absent in 25% of the voiced plosives and suggested that this pattern might be explained by sound change, and might be caused by the large influence of English on Dutch. They also noticed that “there was considerable variation between subjects: some speakers produced prevoicing at the beginning of each voiced plosive, while other speakers only did so for some of the items” (Van Alphen & Smits, 2004: 252). Ziliak and Van de Velde (2008) investigated this phenomenon in the speech of 160 Dutch language teachers, stratified for community (Flanders and the Netherlands) and region (four in each community). In a logatome reading task they found systematic differences between the Flemish and Netherlandic speakers of standard Dutch: devoicing was regularly present in Flanders (in all regions) and only occasionally showed up in the Netherlands. They observed that plosives involved in these devoicing patterns mainly have a shorter duration of prevoicing and a longer duration of the closure. In conclusion, there is acoustic evidence that Dutch initial bilabial stops are sometimes devoiced, especially in Flanders. Therefore, we assume it is an incipient change.

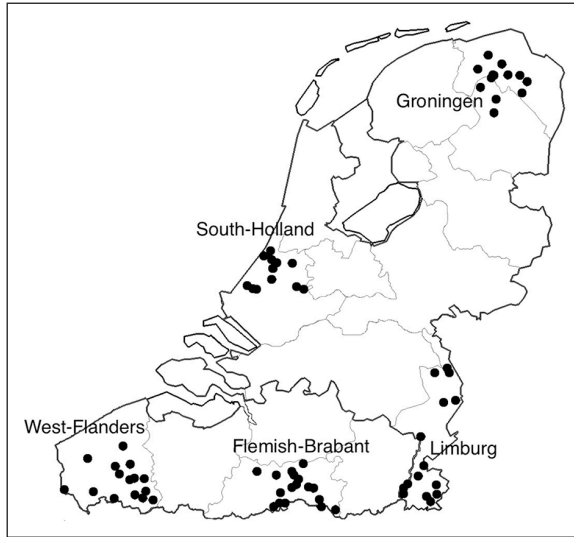


Figure 1. Map of the Dutch language area (the Netherlands and Flanders) and of the five selected regions. Each dot represents the origin of one or more participants ($n = 20$ per region).

The two variables under investigation show similarities, especially from a phonological point of view. They both concern word-initial consonantal realizations in which a voicing contrast is disappearing. However, these changes seem to be the most advanced in different areas (central and northern part of the Netherlands for fricatives, Flanders for stops), and to be unrelated. There are differences in functional load of these contrasts. Dutch only has a few minimal pairs with initial /v/ and /f/ (approximately 10; see CELEX database (Baayen, Piepenbrock, & Gulikers, 1995)) and most of these have a very low frequency in the Dutch language. Dutch bilabial initial stops, in contrast, show many more minimal pairs, which could constitute a blocking or delaying factor for a merger of /b/ and /p/.

2.3 Regions

Five regions within the Dutch language area were chosen: Groningen (GR), South-Holland (SH), Netherlands Limburg (LI), Flemish-Brabant (FB), and West-Flanders (WF) (see Figure 1). Each region belongs to one of the five main dialect groups of Dutch (Taeldeman & Hinskens, 2013: 129–141). Next to covering a broad range of regional variation, this regional approach enables us to reflect different stages of sound change in the two selected variables (see Table 1). The Groningen region (GR) is situated in the north of the Netherlands and is centered around the cities of Groningen and Assen. The region belongs to the Saxonian dialect area. South-Holland (SH) is part of the *Randstad*, the central area in the Netherlands consisting of the urban zone in the western provinces North-Holland, South-Holland, and Utrecht. The chosen region centers around the towns of Leiden and Delft, and belongs to the Hollandic dialect region. Netherlands Limburg (LI) is a geographically peripheral region situated in the south of the Netherlands, stretching from Venlo to Maastricht. It belongs to the Limburgian dialect area. Flemish-Brabant (FB) is the central area in Flanders, having a comparable economic, cultural, and political status in Flanders to South-Holland in the Netherlands. It is situated in the Brabantic dialect area. West-Flanders (WF) is a geographically peripheral region in the western part of Flanders, bordering the North Sea. The

Table I. The five selected regions, their geographical position within the Dutch language area and their phase of the sound change for the two variables under consideration.

Regions	Sources	Stage of devoicing /v/	Stage of devoicing /b/
		Kissine et al., 2003; 2005; Van de Velde, 1996	Van Alphen et al., 2004; Ziliak & Van de Velde, 2008
Groningen (GR)	peripheral region in the Netherlands (NL)	Almost complete	no?
South-Holland (SH)	central area in the Netherlands (NL)	Strong	no?
Limburg (LI)	peripheral region in the Netherlands (NL)	Moderate	no?
Flemish-Brabant (FB)	central area in Belgium (BE)	Weak	Incipient
West-Flanders (WF)	peripheral region in Belgium (BE)	Incipient	Incipient

chosen area is situated around the towns of Kortrijk and Roeselare, and belongs to the Flemish dialect area.

2.4 Participants

The participants were 100 native speakers of Dutch born and raised in these five regions (see Figure 1). Of each region, 10 males and 10 females took part in a series of experiments. The factors *age* and *educational background* were kept constant. The participants were all highly educated young adults aged between 18 and 28 years (mean = 22.03 years). There was no significant difference in age between regions (ANOVA: $df = 4; 95$, $F = 2.161$, $p = .079$). All participants were attending or recently graduated from a university or a university of applied science, and none of them had problems with speaking standard Dutch. Their speech produced during the interview sessions all fits the inclusive definition of standard Dutch (Smakman, 2006; see section 2.1). A total of 36% of the participants reported not to have a regional accent when speaking standard Dutch, 64% reported to have a regionally marked accent. There are regional differences ($df = 4$, $\chi^2 = 15.364$, $p = .004$): the Flemish and Dutch Limburg speakers report most to have an accent (West-Flanders 15/20, Flemish-Brabant 14/20, Limburg 18/20, South-Holland 8/20, Groningen 9/20). No participant reported having any hearing difficulties. In total, 40 participants reported to speak a local dialect in addition to the standard and colloquial varieties of Dutch. The proportion of local dialect speakers differed between regions ($df = 4$, $\chi^2 = 52.083$, $p < .001$) and reflected accurately the current dialect use situation in Flanders and in the Netherlands (West-Flanders 16/20, Flemish-Brabant 4/20, Limburg 17/20, South-Holland 0/20, Groningen 3/20) (Goeman & Jongenburger 2009; Janssens & Marynissen, 2005; Vandekerckhove, 2009).

Both Study 1 on speech production patterns and Study 2 on speech perception were conducted on this pool of participants.

3 Study I: Production

The aim of Study 1 is to investigate the production patterns of the two selected variables in the pool of 100 native speakers.

3.1 Procedure

Participants conducted five different production tasks: word reading, carrier sentence reading, sentence reading, semi-spontaneous speech, and spontaneous speech. All tasks were designed to elicit the standard variety; they thus did not aim at triggering variables in non-standard or local dialect varieties as is usually the case in sociolinguistic studies of language variation and change (see Labov, 1994). This approach has proven to be successful in the study of regional variation in standard Dutch spoken in the Netherlands and Flanders (Adank, Van Hout, & Van de Velde, 2007; Kissine et al., 2003, 2005; Van de Velde & Van Hout, 2002; Van de Velde et al., 2010; Van der Harst, 2011). The participants knew they participated in a study on standard Dutch and the experiment leader (the first author) consistently spoke standard Dutch with all of them. These tasks differ in the amount of attention paid to speech and were intended to elicit the range of phonetic realizations within the standard variety, not to study stylistic differences within the standard variety. All participants conducted the tasks in the same order.

In the word reading task, each participant read a list of words presented in isolation in a randomized order, including words beginning with labiodental voiced fricatives ($n = 20$) and bilabial voiced stops ($n = 20$) and fillers ($n = 82$). In the carrier sentence reading, Dutch non-words beginning with voiced fricatives and stops were produced in the frame “Ik neem de ____” [*I take the ____*] ($n = 9$ per variable). Next, each participant read a set of declarative sentences in which Dutch words starting with voiced fricatives and stops were elicited ($n = 14$ per variable). Semi-spontaneous productions of fricatives and stops were elicited in two pictures-description tasks. The pictures contained a set of objects that the participants were required to name during the description, containing initial fricatives and stops. Spontaneous speech was elicited in an interview carried out by the experiment leader in which participants spoke about some topics related to their daily life. The interview length was approximately 15 minutes. Ten tokens of (v) and of (b) from the semi-spontaneous production task and five tokens of each variable from the semi-spontaneous production task were analyzed, all of them in onset position and preceded by a vowel.

In all five tasks, each participant produced a maximum of 58 (v) tokens and 58 (b) tokens (see Pinget, 2015: 202–208 for a full description of the test materials). All experiments were conducted on a laptop operating with Linux, a Beyerdynamic DT 250 headphone, and an AKG C420 cardioid condenser head-mounted microphone. This equipment was designed for portability, while still providing excellent recordings. Since the same recording and computer equipment was used in the five regions, no apparent difference in the quality of the recorded speech signal and no difference in the subjects' performance related to the testing conditions were observed.

3.2 Phonetic measures

All recordings were sampled at 48 kHz, 24 bits. Realizations of target variables are segmented and labeled into phonetic segments. The segmentation was done by the first author.

3.2.1 Fricatives. Fricatives were segmented by assessing their center of gravity (Gordon, Bartha-maier, & Sands, 2002; Jassem, 1979; van Son & Pols, 1996), following the segmentation protocol for Dutch (Van Son, 2000). The center of gravity (CoG) was calculated in the domain of 0 to 16,000 Hz and the signal was not pre-emphasized prior to weighting. CoG values are characteristically high for fricatives, low for vowels. Since fricatives are characterized by the presence of noise, the onset of fricatives can be determined by the start of noise (rising CoG values) and the offset of fricatives by the end of the noise (falling CoG values). CoG values were determined for all fricatives. Subsequently, onset boundaries were placed manually where the rising slope in CoG started

and the offset boundaries where the falling slope in CoG ended. Voiced fricatives usually show a lower CoG than their voiceless counterparts, since they often have a lower intensity. Despite this difference, boundaries of voiced fricatives were easily detectable with CoG patterns.

Following Kissine et al. (2003), voicing was calculated by measuring the fundamental frequency (F0) (in Hertz) with intervals of 10ms in the segment. The presence of voicing was assessed between 50 and 400 Hz. The number of measurements with presence of F0 was divided by the total number of measurements and multiplied by 100. This resulted in a relative voicing score, indicating the proportion of each fricative segment as voiced. It ranges from 0 (no voicing throughout the fricative) to 100 (voicing throughout the entire fricative).

3.2.2 Stops. Stops were segmented by defining their onset and their offset. The criterion for indicating the onset of the stop is the point where the second formant (F2) disappears, which is equivalent to the offset of the preceding vowel (Cho & McQueen, 2005). The criterion for indicating the offset of the stop is the F2 onset of the next segment, following Rojczyk (2011: 117). The stop onset, offset, and release burst were determined by visual analysis of the spectrographic display and waveforms. There were cases where vocal pulsing dropped before the release or cases where vocal pulsing ceases in the middle of the prevoiced part and started again before the burst (see Pinget, 2015 for more details). Foulkes, Docherty, and Jones (2010) reported this phenomenon in which the cessation of phonation leads to a short period of voicelessness in the prevoicing. Consequently, VOT turned out to be an unreliable measure for voiced stops in this study. This issue led, following Ziliak and Van de Velde (2008), to the introduction of a measure for stops that can capture the gradience of the devoicing phenomenon: voicing. Voicing was measured in stops in the same way as in fricatives, allowing a direct comparison between both variables. Fundamental frequencies (F0) were assessed between 50 and 400 Hz with intervals of 10ms between the onset and offset of the stops. The number of measurements with presence of vocal pulsing was divided by the total number of measurements and multiplied by 100, resulting in a relative voicing index, ranging between 0 (no voicing throughout the stop) and 100 (voicing throughout the whole stop).

3.3 Results

All observations greater than four standard deviations from the mean of the measurements were removed. This conservative procedure managed to remove extremely deviant observations, errors in measurements and mistakes participants made. In total, 5470 voiced fricative and 5576 voiced stop tokens were analyzed as described above. Figure 2 presents boxplots of the results calculated over subject means for each region (for all tasks): (v) voicing measures in the left panel and (b) voicing measures in the right panel.

3.3.1 Fricatives. Voicing measurements in (v) show clear regional differences that are in line with the patterns described in previous studies (Kissine et al., 2003, 2005; Slis & van Heugten, 1989; Van de Velde, 1996) (See Table 1). West-Flemish participants produce (v) with the most voicing, participants from Groningen with the least voicing, and the participants from the other regions have an intermediate position between these extremes. The data are fitted with a mixed-effects linear regression with region as fixed factor; speakers and words as random intercepts and region by words as a random slope. In West-Flanders (the reference level) (v) is most voiced (estimated mean = 65.15%, $t = 10.268$), which does not significantly differ from Flemish-Brabant (estimated mean = 53.53%, $t = -1.574$). Realizations from Limburg (estimated mean = 46.46%, $t = -2.141$), South-Holland (estimated mean = 33.93%, $t = -5.353$) and Groningen (estimated

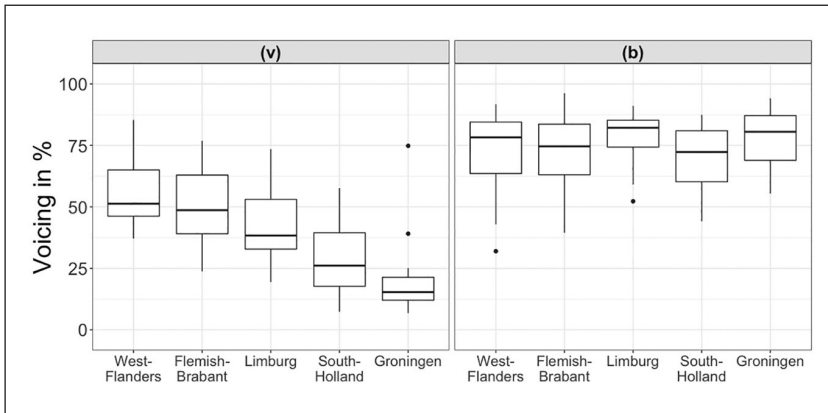


Figure 2. Boxplot of the voicing (in %) per speaker ($N = 100$) of the voiced fricative (in the left panel, $n = 5470$) and voiced stop realizations (in the right panel, $n = 5576$), split up by region.

mean = 25.32%, $t = -5.026$) all significantly differ from realizations in West-Flanders. All pairwise comparisons, except West-Flanders/Flemish-Brabant and Flemish-Brabant/Limburg, are significant.

3.3.2 Stops. Voicing measurements in voiced stops turn out to be consistent across regions. Realizations of (b) are on average highly voiced (around 75–80%). The results are fitted with mixed-effects linear regression with region as fixed factor, speakers and words as random intercepts and region by words as a random slope. The mean voicing for (b) does not show any effect of region: West-Flemish (b) is taken as the reference level (estimated mean = 76.03%, $t = 17.907$), which does not significantly differ from (b) in Flemish-Brabant (estimated mean = 74.633%, $t = -0.316$), Limburg (estimated mean = 80.18%, $t = 0.886$), South-Holland (estimated mean = 73.476%, $t = -0.488$) and Groningen (b) (estimated mean = 79.61%, $t = 0.818$). None of the pairwise comparisons are significant. However, in different regions there are some individuals with clear devoiced realizations of (b). The individual production patterns are linked with the individual perception patterns in section 5 of this paper.

In conclusion, we conducted in Study 1 five different production tasks designed to elicit realizations labiodental fricatives and bilabial stops in the standard variety. We found clear regional differences in production of the fricative contrast, which are in line with previous research on this sound change in progress (see section 2.3.). The production patterns of bilabial stops did not show any region effects; devoicing can only be seen at the individual level.

4 Study 2: Perception

The aim of Study 2 is to investigate the perceptual patterns of the two selected variables in the pool of 100 native participants. An analysis of the regional differences in the perception of fricatives was previously reported in Pinget et al. (2016).

4.1 Stimuli

The fricative and stop stimuli used in Study 2 are described here.

4.1.1 Fricatives. A speech continuum between /v/ and /f/ was generated by manipulating the proportion of voicing. Fricatives were presented in a CV syllable. As /i/ is the vowel showing the least regional variation in standard Dutch (Van der Harst, 2011: 159), it was chosen to avoid a bias caused by regional differences in the perception of the vowel. Naturally produced syllables /vi/ and /fi/ produced by a male native speaker of Dutch (25 years old, from the South-Holland region, trained phonetician) were digitally recorded with a sample frequency of 44.1 kHz in a sound-attenuated cabin. The fricatives of the source recordings were extracted from their original context and used as the extremes of the continuum along the voicing dimension. The nine steps along the first dimension were generated by spectral linear interpolation, using the PSOLA (Pitch-Synchronous-Linear-Overlap-and-Add) algorithm of Praat (based on the script of Mitterer, 2009; Boersma & Weenink, 2014). Besides the two extremes of the continuum with respectively 0% and 100% voicing, the interpolation provided seven realizations characterized by approximately 12.5%, 25%, 37.5%, 50%, 62.5%, 75%, and 87.5% voicing. The nine steps of the continuum were originally also manipulated for duration, but the effect of duration did not turn out to be significant (see Pinget, 2015; Pinget et al., 2016) and will not be discussed in this paper. The syllable stimuli were obtained by concatenating the /v/-/f/ realizations with the [i] produced in the original /vi/ context. The vowel had a constant duration of 110ms and its pitch contour was manipulated through the PSOLA pitch manipulation and LPC resynthesis functions of Praat. In the vowel transition, the pitch contour was flattened to 135 Hz. In order to produce a falling contour, the pitch contour was gradually reduced to 120 Hz from 60ms after the start of the vowel until its endpoint.

4.1.2 Stops. A speech continuum between /b/ and /p/ was generated by manipulating VOT. The manipulation provided nine intermediate realizations characterized by a VOT of -90ms, -74ms, -58ms, -42ms, -26ms, -10ms, 6ms, 22ms, and 38ms. These duration endpoints correspond to the average VOTs reported by Van Alphen and Smits (2004). To manipulate VOT, parts of the prevoicing were gradually removed from the original voiced stop realizations in order to shorten the negative VOT. For the steps with positive VOT, parts of the burst from the voiceless stops were added resulting in extended positive VOT. The nine steps of the continuum were originally also manipulated for duration, but the effect of duration did not turn out to be significant (see Pinget, 2015; Pinget et al., 2016) and will not be discussed in this paper. It will therefore be left out of consideration here. The same PSOLA pitch manipulation as for fricatives was applied. Stops were presented in a CV syllable (with /i/).

4.2 Procedure

Participants took part in two speeded forced-choice identification tasks: one with fricative stimuli and one with stop stimuli. They listened in a sound-attenuated booth to the stimuli through a headphone Beyerdynamic DT 250. They were asked to categorize the realizations as being /f/ or /v/ (or /p/-/b/) sound by pressing the red or blue button of a button box labeled with the corresponding sounds. The order of presentation of the consonants on the button box was balanced between participants. The task was self-paced. Reaction time (RT) was recorded from the beginning of the stimuli. Participants had a time window of 800ms after the end of the stimulus to give their response. A response given after this time window was discarded to avoid responses resulting from second guesses. Twelve stimuli were presented in the practice session to familiarize the participant with the task. The task consisted of five blocks of 81 trials each (9*9) in which the stimuli were presented randomly (i.e., 405 trials per participant). Trials with response times shorter than the duration of the consonants were excluded from the analysis, as these responses were given before the participant could hear the entire first segment. The binary responses obtained in this speeded

forced-choice identification tasks were analysed using logistic regressions, with the identification as /v/ (or /b/) as the dependent variable. In logistic regressions the probability of x ($P(x)$) – in this case the probability of a /v/ or /b/ response – is predicted by the following equation:

$$P(x) = e(\beta_0 + \beta_1 x) / (1 + e(\beta_0 + \beta_1 x))$$

β_0 is the estimate of the intercept and β_1 is the estimate of the slope of the logistic regression line. The higher the absolute value of β_1 , the steeper the slope of the regression line. In order to obtain in the logistic regressions parameters that are easily interpretable, the continuum steps were centralized with steps ranging from -4 to 4. Logistic regressions provide identification curves that are defined by the slopes, which indicate how categorical the binary judgment is (the steeper the curve, the more categorical the judgment), and the cut-off point between the voiced and voiceless categories: i.e., the point where the probability $P(x)$ is equal to .5). The cut-off point between the voiced and voiceless categories can be calculated on the basis of these two estimates with the following formula (Kendall & Fridland, 2012):

$$\text{Cut-off point} = -\beta_0 / \beta_1$$

4.3 Results

Excluded data (too short latency) and missing values (too long latency) represent 6.97% of the fricatives data and 4.21% of the stops data. In total, 37,678 valid fricative observations and 38,803 valid stop observations were used for the quantitative analysis. The results for both fricatives and stops are shown in Figure 3.

4.3.1 Fricatives. A mixed model logistic regression with region and voicing steps along the continuum as fixed factors, listeners and trials as random intercepts and voicing steps by listener as random slopes was fitted to the binary response data. The model—which corroborates the analyses provided in Pinget et al. (2016)—shows that there is a significant effect of voicing steps. This slope is negative ($\beta = -1.055$, $p < .001$): the more to the left (the more voicing in the acoustic signal), the more /v/ responses. The Flemish regions West-Flanders and Flemish-Brabant have the steepest slope along the voicing dimension (respectively $\beta = -1.095$ and $\beta = -0.956$). The Dutch regions: Limburg ($\beta = -0.835$), South-Holland ($\beta = -0.761$) and Groningen ($\beta = -0.672$) have less steep slopes than the Flemish regions. West-Flemish listeners almost categorically respond /v/ in the first three steps of the continuum (-4 to -2), while participants from other regions on average start lower and gradually decrease along the continuum. The 0.5 cut-off point was calculated, following Kendall and Fridland (2012), as the median (i.e., the point where the probability $P(x)$ is equal to .5). In all regions, the cut-off point was situated to the right of our continuum (between centralized step +1 and +2), but does not significantly differ across regions (ANOVA: $F < 1$).

4.3.2 Stops. The perception of bilabial stops is highly categorical, since most stimuli on the continuum were clearly identified either as /b/ or /p/. Only for steps +0, +1 and +2, the decisions were more variable and the cut-off point between the categories lies around step +1. The mixed model logistic regression with region and VOT steps along the continuum as fixed factors, listeners and trials as random intercepts, and voicing steps by listener and region by trial as random slopes

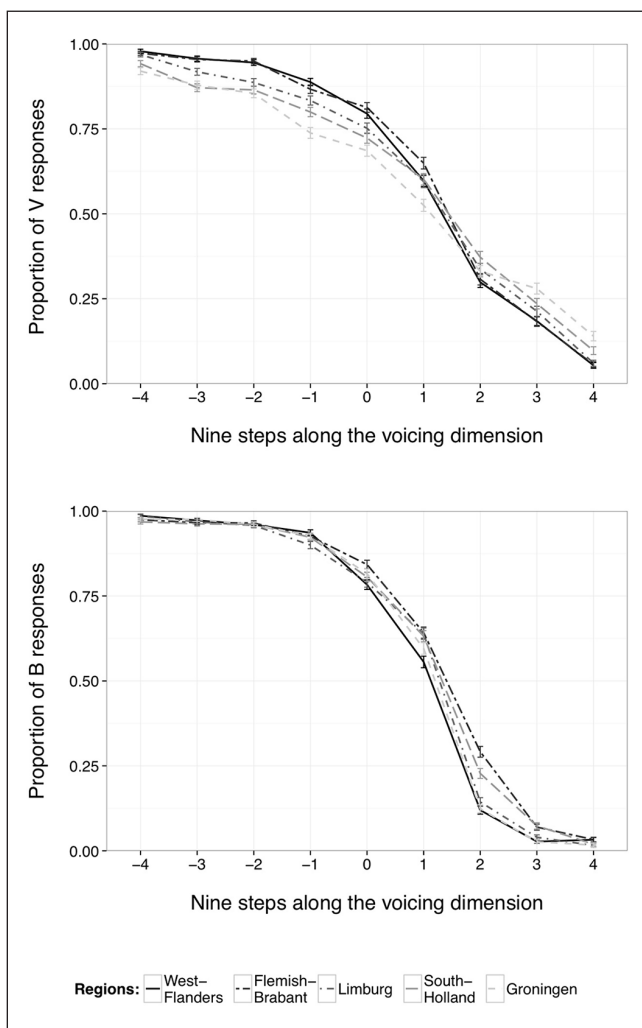


Figure 3. Results of the identification task along the voicing/VOT dimension, split up by region. The results for fricatives are presented in the upper panel and for stops in the lower panel. The centralized nine-steps continuum is presented on the x-axis. The leftmost part (negative values) refers to the most voiced realizations and the rightmost part (positive values) refers to the voiceless realizations. The proportion of voiced (/v/ or /b/) responses is presented on the y-axis. Error bars represent ± 1 standard error.

shows that participants from the five regions have a highly similar perception of the /b/-/p/ contrast. There is a significant effect of VOT steps: the slope is negative ($\beta = -1.420, p < .001$), indicating that the more to the left (the more negative the VOT), the more /b/ responses are given. There is, however, no significant difference between regions (all z values < 1). In conclusion, participants from all regions show highly similar perceptual patterns along the VOT dimension.

In conclusion, we conducted in Study 2 two speeded forced-choice identification tasks (one with labiodental fricative stimuli and one with bilabial stop stimuli) to test the perception of these contrasts. We found at the group level that regional differences occur in the speech perception of

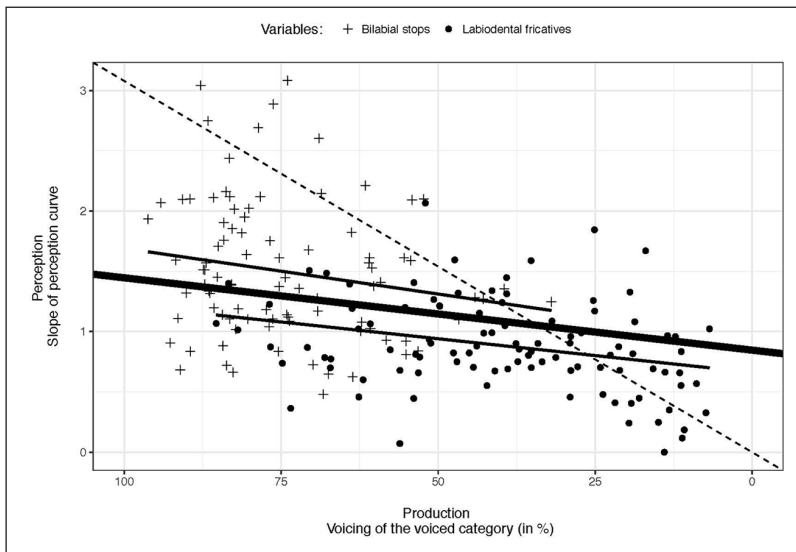


Figure 4. Scatterplot of the slope of the individual perception curves (y-axis) and the voicing of the voiced category (x-axis, range: 100 (fully voiced, no change yet) to 0 (fully voiceless, complete change)), split up by variable (dots for labiodental fricatives and pluses for bilabial stops). The thick solid line represents the fitted line to all data. The two thin solid lines represent the fitted lines to the data for fricatives (lower thin line) and stops (upper thin line) separately. The dashed line represents the hypothetical fitted line in case of a perfect correlation between production and perception.

labiodental fricatives in the Dutch language. These regional differences relate to the perception slopes and not to the cut-off point between categories. For bilabial stops however, we did not find any regional differences in the perception of the /p-/b/ contrast.

5 Link between variation in production and perception

In order to obtain insight into the relationship between perception and production at the individual level, the production measures from Study 1 and the perception measures from Study 2 are now combined.

5.1 Relationship between individuals

The production measure from Study 1 (i.e., the voicing of (v) and (b), ranging between 0 and 100 %) turned out to capture the phases of sound change in production (see Figure 2). The perception measure from Study 2 (i.e., the slope of categorical perception) was measured as the slope of the logistic regression curves obtained in the identification experiment (see Figure 3). The perception slope was calculated for each individual listener in a separate logistic regression model with voicing steps along the continuum as fixed factor and trials as random intercept. The slopes obtained in this manner correlated strongly with the slope measures obtained from the random effect structure of the model presented in section 4.3 ($r = .95, p < .001$). The individual perception slopes ranged between 0 (no categorical perception at all) and 3.083 (strong categorical perception).

In Figure 4, the individual measures of speech production and perception are correlated. The x-axis shows the voicing of the voiced category (in %). The axis has been flipped around to

represent the course of sound change: from 100% (beginning of the sound change) to 0% (end of the sound change). The y-axis shows the individual slope of the perception curve. Plusses represent individual results for bilabial stops and dots for labiodental fricatives. There are four fitted lines: the dashed line shows the (hypothetical) case of a perfect correlation between production and perception, the lower thin line represents the fricatives, the upper thin line the stops, and the thick solid line all data.

The following can be observed in Figure 4. First, the slope in perception shows a relationship with the voicing of the voiced category in production. The slope in perception decreases in a linear manner as the change is proceeding and the voiced category is devoicing (thick solid line). The slope in perception is moderately predicted by the voicing of the voiced category in a mixed-effects linear regression (with voicing, variable, and their interaction as fixed factors; speakers as random intercept ($t = 2.096$, Marginal $r^2_{\text{GLMM}} = .321$, Conditional $r^2_{\text{GLMM}} = .425$); see Nakagawa, Johnson, & Schielzeth, 2017 for more explanation on the calculation of r^2 for mixed-effects regressions). The more advanced the change, the lower the perception slope, thus the less categorical the perception. The pattern is present for both fricatives (lower thin line) and stops (upper thin line). This relationship constitutes evidence for a link between speech production and speech perception.

Second, the range of perception slopes (visualized on the vertical axis of Figure 4) in stops (plusses) is larger than in fricatives (dots). Larger variation in perception is expected at the beginning of a sound change. Near the completion of sound change in progress, it appears that the values of perception slopes are in most cases not reaching 0. So, there is still some form of categorical perception in most individuals who are advanced in the merging process in production. Listeners who merge /v/ and /f/ mostly in production are still able to distinguish /v/ and /f/, as can be seen in the right part of Figure 4 and in the fact that the fitted line for fricatives does not decrease steeply.

Finally, we want to draw the attention to the dashed line representing the relationship between perception and production measures in the case of a simultaneous change of production and perception. This equation line has a zero y-intercept (i.e., 0% of voicing in production and no perception slope reflect the typical end stadium of this particular change) and crosses the y-axis at a slope of value of 3.083 (i.e., the slope value for the most categorical individual perception in our data). This dashed line is based on the assumption that a 3.083 slope value corresponds to maximally categorical perception, but this slope value just happens to be the steepest slope in this sample. We are aware that the upper limit of categorical perception in a larger/different sample might have been higher. This analysis however allows a conservative comparability between perception and production at the individual level, as it will be developed in the next section.

5.2 Relationship within individuals

The current analysis focuses on the relationship between speech production and speech perception at the individual level. In the previous section, we investigated this relationship between individuals. In this section however, we investigate the relationship production-perception within individuals. Therefore, a measure derived from the previous measures presented in section 5.1. was computed for each individual: the difference production-perception (diffPP). It is calculated as follows:

$$\text{diffPP} = \text{difference in voicing between categories} - \text{slope in perception} * 32.436$$

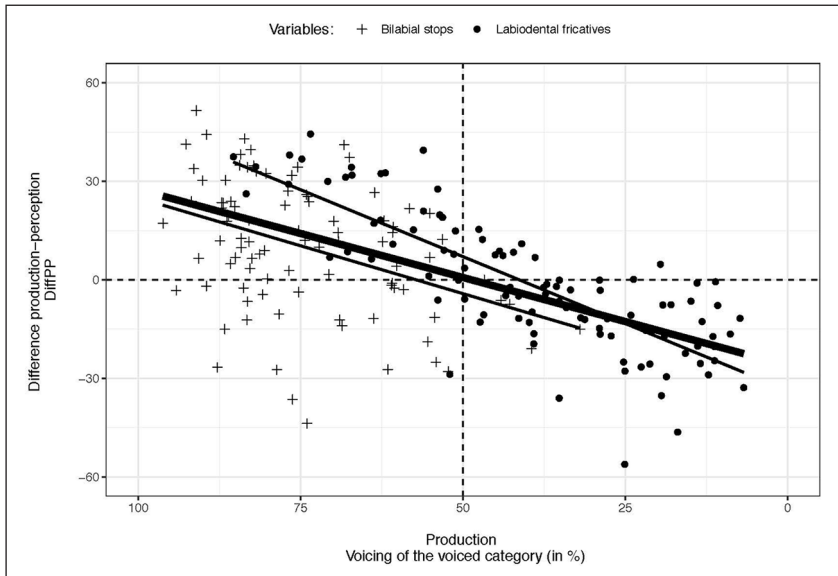


Figure 5. Scatterplot of the production-perception difference (diffPP), as a function of the voicing of the voiced category (ranged from 100 (no change yet) to 0 (fully completed devoicing category)), split up by variable (dots for labiodental fricatives and plusses for bilabial stops). The thick solid line represents the fitted line to all data. The two thin solid lines represent the fitted lines to the data of the fricatives (lower thin line) and the stops (upper thin line) separately.

As the most categorical slope in the sample had a value of 3.083, the perception slopes are multiplied by 32.436 ($100/3.083$), to obtain a maximal perception score of 100 (maximally categorical). The diffPP per individual is computed as the difference between—on the one hand—the difference in voicing between the voiced and voiceless categories in production and—on the other hand—the perception score. The calculation of the diffPP departs from the underlying assumption that both production and perception can be measured on the same scale, namely ranging from 0 (no difference in production between the two categories and perception of one category only) to 100 (maximum possible difference in production between the two categories and maximally categorical perception). The diffPP constitutes a measure of the relationship between an individual's state of speech production and speech perception. It is the distance between individual observations and the regression line of a hypothetical simultaneous change plotted as a dashed line in Figure 4. If the measure is positive, an individual's production score is higher than the perception score and the change is less advanced in production than in perception. If the measure equals 0, the individual's production and perception are equally advanced in the process of change. If the measure is negative, the individual perception score is higher than the production score and the change is less advanced in perception than in production.

In Figure 5, diffPP (y-axis) is plotted against the voicing of the voiced category (x-axis, range: 100 (fully voiced, no change yet) to 0 (fully voiceless, complete change)). The data are split up by variable (dots for labiodental fricatives and plusses for bilabial stops). There are three fitted lines: the thick solid represents all data, the lower thin line the fricatives and the upper thin line the stops.

At the beginning of the change, most participants show a positive diffPP, indicating that the change is more advanced in their perception than in their production. However, some participants show—already at the beginning of the change—a slightly negative diffPP, indicating that the

change is a bit more advanced in their production than in their perception (i.e., observations in the lower left quarter in Figure 5). At around 50% of voicing (individuals half-way the change in production), the diffPP often equals 0 with individual production and perception approximately at the same stage of advancement. Towards the end of the sound change, the diffPP is mostly negative, indicating that the change is more advanced in production than in perception. The variation in stops is larger than in fricatives, which is consistent with the fact that stop devoicing is still an incipient change. This relationship is tested with a linear regression analysis.

The diffPP is significantly predicted by the voicing of the voicing category in a mixed-effects linear regression (with voicing, variable, and their interaction as fixed factors; speakers as random intercept ($t = 5.295$, Marginal $r^2_{\text{GLMM}} = .461$, Conditional $r^2_{\text{GLMM}} = .582$); see Nakagawa et al., 2017 for more explanation on the calculation of r^2 for mixed-effects regressions). The fact that this relationship is statistically significant is not directly informative, since the computation of the diffPP is partly based on the voicing of the voiced category. However, the fact that this relationship is negative is of particular interest for the order of change in production and perception systems. Indeed, if change in production and perception were to happen simultaneously, participants would have a diffPP around 0 all the way through the change, which is definitely not the case here. If change in perception would be completed before change in production, the fitted slope would be positive instead of negative. However, the slope is negative. When a change is nearing completion, individual diffPP values are below 0, indicating that production is further advanced than perception: the voiced category is reaching complete devoicing, whereas in perception voiced and voiceless categories are still distinguished. Hence it constitutes evidence that in the earlier stages perception changes before production, but that towards the end of the sound change perception lags behind.

It should be noted that this analysis within individuals is based on the assumption that one of the participants exhibited maximally categorical perception. As explained in section 5.1, participants of a larger/different sample might have exhibited an even stronger categorical perception. This argument does not undermine our conclusions, since the patterns described here would be even stronger with a higher maximal categorical perception (i.e., perception would be even more advanced than production at the beginning of the change, and production would be even more advanced than perception at the end of the change). The current analysis thus constitutes a rather conservative way to compare perception and production.

6 Discussion

The two studies were designed to investigate how patterns of variation and change in speech production relate to speech perception, and whether change in speech perception precedes or follows change in speech production. The analyses were based on the analysis of two changes, one nearly completed and the other incipient. The juxtaposition of these two consonantal changes in progress allows one to obtain insights into the whole process of sound change. The results of these two studies provide three types of evidence in favor of a link between speech perception and speech production in the context of sound change: (a) an analysis at the regional level, (b) an analysis between individuals, and (c) an analysis within individuals.

First, it was shown at the group level that regional differences occur in the speech perception of labiodental fricatives in the Dutch language (see section 4). These regional differences relate to the perception slopes and not to the cut-off point between categories. This means that there was no difference with respect to the category boundary, but only with respect to the degree of categoricalness in the identification. Crucially, the differences in perception between regions match production patterns reported in section 3 and in previous studies (Kissine et al., 2003; 2005; Van Alphen

et al., 2004; Van de Velde, 1996; Ziliak & Van de Velde, 2008). The less devoicing in speech production in a region, the more categorical the stimuli are perceived by participants from this region. These results at the group level confirmed for consonants what Fridland and Kendall (2012) and Kendall and Fridland (2012) found for vowels.

Second, the correlation between speech perception and speech production at the individual level turned out to be moderate (Marginal $r^2_{\text{GLMM}} = .321$, Conditional $r^2_{\text{GLMM}} = .425$). Because of the type of regression, the strength of the relationship between production and perception in our study is difficult to compare with previous studies. To recall, Newman (2003) found significant correlations between listeners' perceptual prototypes and their average VOTs for English stop consonants ($r = .49, p = .05$). Beddor and Coetzee (2014) found a weak correlation ($r = .27, p = .01$) between production and perception.

Third, the analysis within individuals suggests that most individuals who start to participate in a sound change, change their perceptual patterns first. These perceptual adjustments are then followed by changes in their speech production patterns. Once the devoicing is launched in production, the endpoint appears to be the merger of the voiced and voiceless category. However, it is shown that perception is not as close to this endpoint than production: perception lags behind in the final phases of sound change. Participants are still able to hear some categorical contrast in a variable for which they do not make the contrast in production themselves. The fact that the sound change seems to be triggered by a change in perception first, which is followed by a change in production, is easily conceivable. As the ambient forms reach the listeners through the perceptual system, it seems logic that a change happens first in this system. At the same time, it is plausible that change in perception does not reach an end stage as easily as change in production. Indeed, listeners still (need to) maintain some perceptual contrast caused by and to cope with speech input they receive from speakers from outside their region or from other age groups (in which the contrast is still present) in both mass media and personal communication. It should be noted that the Dutch language area is only about 55,000 km², that the area is densely populated, and that mobility is high. Moreover, the fact that the Dutch society has a high degree of literacy and exposure to written media (the voicing differences are present in orthography) presumably contributes to a remaining awareness of these variables.

Even if the two sound changes in this study are very similar types of change, it is legitimate to ask whether their intrinsic differences do not bias the results. As far as speech perception is concerned, it has for example been proposed that the perception of stops is intrinsically more categorical than the perception of fricatives (Healy & Repp, 1982; Liberman et al., 1967). In the literature, evidence is conflicting. Although it is generally accepted that overall the perception of consonants is more categorical than the perception of vowels (Fry et al., 1962; Pisoni, 1975; Repp, 1984), there is some disagreement over the degree of categorical perception within consonants. Repp (1981) showed that fricatives follow patterns similar to the categorical perception found in stop consonants. In contrast, other studies found that fricative perception is less categorical than stop perception (Healy & Repp, 1982; Liberman et al., 1967). Looking at the current perception study (in section 4), it is clear that the perception of stops shows broader variation than the perception of fricatives. Moreover, the conclusions reached in section 5 still hold when the stop data are left out of consideration. The negative correlations in both Figure 4 and Figure 5 are also significant on the fricative data only (Linear Regression in Figure 4 for fricative observations only: $t = 3.036, p = .003, r^2 = .086$; and Linear Regression in Figure 5 for fricative observations only: $t = 13.090, p < .001, r^2 = .636$). Based on these arguments, it is concluded that intrinsic differences in the perception of stops and fricatives (if they exist) do not undermine our conclusions. As far as speech production is concerned, there also might be intrinsic differences between the process of devoicing in stops and in fricatives. Fricative devoicing might be more likely to occur due to the difficulty to

maintain voicing in turbulent sounds (i.e., to overcome the conflict between high supralaryngeal air pressure needed for fricative versus lower supralaryngeal air pressure that is necessary to preserve vocal fold vibration). Moreover, the functional load of the fricative and stop contrasts is not identical: there are only a few minimal pairs with initial /v/ and /f/, whereas bilabial initial stops show many more minimal pairs, which could constitute a blocking or delaying factor for the devoicing of /b/.

It should also be noted that the measure of relationship between individual's production and perception (diffPP) constitutes indirect evidence for the order in which perception and production are changing within an individual: it gives a state of the relationship at one specific point in time. Only a longitudinal study of individuals' production-perception link would provide direct evidence. Nevertheless, this measure showed that change in perception seems to precede change in production and that perception lags behind at the end of the sound change. These data support the view of Harrington et al. (2008) that there is not always a uniform relationship between the phonological system and phonetic output for all members of the same speech community. Instead, phonological categories are likely to be related to the differences in individual production.

Finally, it is necessary to raise the question whether the second variable in the study, the devoicing of bilabial stops, is really a sound change in progress. The devoicing of bilabial stops was chosen to represent an incipient change in the Dutch language. Based on the results of the study of Ziliak and Van de Velde (2008) differences between Flemish and Netherlandic speakers of Dutch in the realization of stops were expected. However, in this study, individual speakers with devoiced stops showed up in all regions and there was no sign of rapid rise in one of the regions. Therefore, it is not clear (yet) whether the devoicing of stops is a sound change in an incipient stage or an idiosyncratic, stable variation pattern, but this does not undermine our interpretation. In the framework of language variation and change, Chambers (2002: 361) has distinguished three phases in the process of the spread of language change: initial stasis, rapid rise and tailing off. Every change starts with a period of initial stability, and the exact turning point between initial stasis and acceleration is not only difficult to determine, but also appears to be different for every change. This view is also at the core of Ohala's work: changes are drawn from the pool of existing variation when some formerly stable variant takes on social meaning (Ohala, 1981). As a result, it is impossible to predict whether a turning point will appear for Dutch stops and what its exact timing will be. Only the future will tell whether and when stop devoicing in Dutch will continue.

The results of this study clearly showed that the relation between perceived and produced speech at the level of the individual is a central issue for the study of the phonetic basis of sound change. The fact that speech perception seemed to be ahead of production in the beginning of the change relates to approaches of sound change in which it is claimed that variants that are somehow salient in perception will be subsequently realized in the speaker's own productions (Beddor, 2012; Harrington et al., 2008; Lindblom, Guion, Hura, Moon, & Willerman, 1995; Ohala, 1981). These initial phases in the sound change can also be linked to the concept of *near-merger* as defined by Labov (1994) and Labov, Karen, and Miller (1991) where a speaker consistently makes a small articulatory difference between two categories, but cannot distinguish these perceptually.

7 Conclusion

Although different views exist on the relation between the processes of perception and production, there is still not a great deal of direct, empirical evidence bearing on the existence of a production-perception link. For the most part, speech perception and speech production have been investigated independently. This paper breaks with this research practice and reports on two studies conducted

on the same pool of 100 participants stratified for region, where the question of the existence of a production-perception link is tackled in the context of sound change. The juxtaposition of two consonantal changes in progress allows us to obtain insights into the whole process of sound change (from its earliest beginning to the end point) and points to effects related to different phases of the process. The analyses of variation in both production (Study 1) and perception (Study 2) on the regional and individual level show that there is a clear relationship between production and perception in sound change. Moreover, it is shown that changes in speech perception precede changes in production at the beginning of sound change, but that speech perception lags behind speech production when the sound change is reaching completion.

Acknowledgements


We are very grateful to the three anonymous reviewers for helping to clarify and improve this work. We thank the following institutes for providing research facilities: the department of Dutch Linguistics of Ghent University, the Leiden University Phonetics Laboratory and the Centre for Language and Cognition Groningen of the University of Groningen, the department of Psychology of the Vrije Universiteit Brussels and the department of Linguistics of Radboud University Nijmegen. We are also very grateful to Hans Rutger Bosker, whose voice was used for the construction of the stimuli, Theo Veenker for his technical assistance in programming, and Mattis van den Bergh and Hugo Quené for their statistical advice.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Netherlands Organization for Scientific Research (NWO grant 322-75-002).

ORCID iDs

Anne-France Pinget  <https://orcid.org/0000-0002-8217-0032>

Hans Van de Velde  <https://orcid.org/0000-0003-2197-5555>

References

- Adank, P., van Hout, R., & Van de Velde, H. (2007). An acoustic description of the vowels of Northern and Southern standard Dutch II: Regional varieties. *The Journal of the Acoustical Society of America*, 121, 1130–1141.
- Ainsworth, W., & Paliwal, K. (1984). Correlation between the production and perception of the English glides /w, r, l, j/. *Journal of Phonetics*, 12(3), 237–243.
- Auer, P. (2005). Europe's sociolinguistic unity, or: A typology of European dialect-standard continua. In N. Delbecque, J. Van der Auwera, & D. Geeraerts (Eds.), *Perspectives on variation: Sociolinguistic, historical, comparative* (pp. 7–42). Berlin, Germany; New York, NY: Mouton de Gruyter.
- Baayen, R., Piepenbrock, R., & Gulikers, L. (1995). *CELEX2 LDC96L14*. Web Download. Philadelphia, PA: Linguistic Data Consortium.
- Bailey, P. J., & Haggard, M. P. (1973). Perception and production: Some correlations on voicing of an initial stop. *Language and Speech*, 16(3), 189–195.
- Beddor, P. (2012). Perception grammars and sound change. In M.-J. Solé & D. Recasens (Eds.), *The initiation of sound change: Production, perception, and social factors* (pp. 37–55). Amsterdam, Netherlands: John Benjamins.
- Beddor, P., & Coetzee, A. (2014). *Sound change propagation: The relation between perception and production in individual language*. Presentation at the workshop Sound Change in Interacting Human Systems. Berkeley, CA.
- Bell-Berti, F., Raphael, L. J., Pisoni, D. B., & Sawusch, J. R. (1979). Some relationships between speech production and perception. *Phonetica*, 36(6), 373–383.

- Boersma, P., & Weenink, D. (2014). *Praat: Doing Phonetics by Computer* [Computer software]. Version 5.3. 84.
- Booij, G. E. (1995). *The phonology of Dutch*. Oxford, UK: Clarendon Press.
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. I. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, 101(4), 2299–2310.
- Cassier, L., & Van de Craen, P. (1986). Vijftig jaar evolutie van het Nederlands [Fifty years evolution of the Dutch language]. In J. Creten, G. Geerts, & K. Jaspaert (Eds.), *Momentopnamen van de sociolinguïstiek in België en Nederland* (pp. 59–73). Leuven, Belgium: Acco.
- Chambers, J. (2002). Patterns of variation including change. In *The handbook of language variation and change* (pp. 349–372). Oxford, UK: Blackwell Publishing Ltd.
- Cho, T., & McQueen, J. M. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, 33(2), 121–157.
- Cohen, A., Ebeling, C. L., Fokkema, K., & Holk, A. G. F. (1961). *Fonologie van het Nederlands en het Fries* [Phonology of Dutch and Frisian]. Leiden, Netherlands: Martinus Nijhoff.
- Debrock, M. (1977). An acoustic correlate of the force of articulation. *Journal of Phonetics*, 5, 61–80.
- Donaldson, B. (1983). *Dutch: A linguistic history of Holland and Belgium* (pp. 3–183). Leiden, Netherlands: Martinus Nijhoff.
- Evans, B. G., & Iverson, P. (2004). Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences. *The Journal of the Acoustical Society of America*, 115(1), 352–361.
- Evans, B. G., & Iverson, P. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *The Journal of the Acoustical Society of America*, 121(6), 3814–3826.
- Ferguson, S., & Quené, H. (2014). Acoustic correlates of vowel intelligibility in clear and conversational speech for young normal-hearing and elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 135(6), 3570–3584.
- Foulkes, P., Docherty, G., & Jones, M. (2010). Analysing stops. In M. Di Paolo & M. Yaeger-Dror (Eds.), *Sociophonetics* (pp. 58–71). Abingdon, UK: Routledge.
- Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, 14(1), 3–28.
- Fox, R. A. (1982). Individual variation in the perception of vowels: Implications for a perception-production link. *Phonetica*, 39(1), 1–22.
- Fridland, V., & Kendall, T. (2012). Exploring the relationship between production and perception in the mid front vowels of US English. *Lingua*, 122(7), 779–793.
- Frieda, E. M., Walley, A. C., Flege, J. E., & Sloane, M. E. (2000). Adults' perception and production of the English vowel /i/. *Journal of Speech, Language, and Hearing Research*, 43(1), 129–143.
- Fry, D. B., Abramson, A. S., Eimas, P. D., & Liberman, A. M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech*, 5(4), 171–189.
- Garrett, A., & Johnson, K. (2013). Phonetic bias in sound change. In A. C. L. Yu (Ed.), *Origins of sound change: Approaches to phonologization* (pp. 51–97). Oxford, UK: Oxford University Press.
- Geeraerts, D., Grondelaers, S., & Speelman, D. (1999). *Convergentie en divergentie in de Nederlandse woordenschat. Een onderzoek naar kleding- en voetbalterminen*. [Convergence and divergence in the Dutch lexicon. A study of clothing and soccer terminology]. Amsterdam, Netherlands: Meertens Instituut.
- Geeraerts, D., & Van de Velde, H. (2013). Supra-regional characteristics of colloquial Dutch. In F. Hinskens & J. Taeldeman (Eds.), *Language and Space* (pp. 532–556). Berlin, Germany: Mouton de Gruyter.
- Goeman, T., & Jongenburger, W. (2009). Dimensions and determinants of dialect use in the Netherlands at the individual and regional levels at the end of the twentieth century. *International Journal of the Sociology of Language*, 196–197, 31–72.
- Goldinger, S. D. (1997). Speech perception and production in an episodic lexicon. In K. Johnson & J. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 33–66). New York, NY: Academic Press.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251.

- Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. In N.O. Schiller & A.S. Meyer (eds.), *Phonetics and Phonology in Language Comprehension and Production*, pp. 159–207. Mouton de Gruyter.
- Grosvald, M. (2009). *Long-distance coarticulation: A production and perception study of English and American Sign Language*. PhD Dissertation. University of California Davis, USA.
- Grosvald, M., & Corina, D. (2012). The production and perception of subphonemic vowel contrasts and the role of the listener in sound change. In M. Solé & D. Recasens (Eds.) *The initiation of sound change: Perception, production, and social factors* (pp. 77–100). Amsterdam, Netherlands: John Benjamins Publishing.
- Gordon, M., Barthmaier, P., & Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association*, 32, 141–174.
- Goossens, J. (1974). *Historische Phonologie des Niederländisches* [Historical Phonology of the Netherlands]. Tübingen, Germany: Niemeyer.
- Gussenhoven, C. (1999). Illustrations of the IPA: Dutch. In *Handbook of the International Phonetic Association* (pp. 74–77). Cambridge, UK: Cambridge University Press.
- Gussenhoven, C., & Bremmer, R. H. (1983). Voiced fricatives in Dutch: Sources and present-day usage. *North-Western European Language Evolution*, 2, 55–71.
- Hamann, S., & Sennema, A. (2005). Acoustic differences between German and Dutch labiodentals. In *ZAS papers in Linguistics*, 42, 33–41.
- Harrington, J., Kleber, F., & Reubold, U. (2008). Compensation for coarticulation: An acoustic and perceptual study. *The Journal of the Acoustical Society of America*, 123(5), 2825–2835.
- Harrington, J., Kleber, F., & Reubold, U. (2012). The production and perception of coarticulation in two types of sound changes in progress. In S. Fuchs, M. Weirich, P. Perrier & D. Pape (eds.) *Speech Planning and Dynamics* (pp. 39–62) Frankfurt: Peter Lang.
- Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics*, 31(3), 373–405.
- Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34(4), 458–484.
- Healy, A. F., & Repp, B. H. (1982). Context independence and phonetic mediation in categorical perception. *Journal of Experimental Psychology: Human Perception and Performance*, 8(1), 68–80.
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). Codes and their vicissitudes. *Behavioral and Brain Sciences*, 24(5), 910–926.
- Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience*, 7(12), 942–951.
- Janson, T. (1983). Sound change in perception and production. *Language*, 59(1), 18–34.
- Janssens, G., & Marynissen, A. (2005). *Het Nederlands vroeger en nu* [Dutch from the past and Dutch now] (pp. 1–263). Leuven, Belgium: Acco.
- Jassem, W. (1979). Classification of fricative spectra using statistical discriminant functions. In B. Lindblom & S. Öhman (Eds.), *Frontiers of speech communication research* (pp. 77–91). New York, NY: Academic Press.
- Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J.W. Mullennix (eds.) *Talker Variability in Speech Processing* (pp. 145–165). San Diego: Academic Press.
- Johnson, K., Flemming, E., & Wright, R. (1993). The hyperspace effect: Phonetic targets are hyperarticulated. *Language*, 69(3), 505–528.
- Keating, P. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, 60(2), 286–319.
- Kendall, T., & Fridland, V. (2012). Variation in perception and production of mid front vowels in the US Southern Vowel Shift. *Journal of Phonetics*, 40(2), 289–306.
- Kissine, M., Van de Velde, H., & van Hout, R. (2003). An acoustic study of standard Dutch /v/, /f/, /z/ and /s/. *Linguistics in the Netherlands*, 20(1), 93–104.

- Kissine, M., Van de Velde, H., & van Hout, R. (2005). Acoustic contributions to sociolinguistics: Devoicing of /v/ and /z/ in Dutch. *University of Pennsylvania Working Papers in Linguistics*, 10(2).
- Kleber, F., Harrington, J., & Reubold, U. (2012). The relationship between the perception and production of coarticulation during a sound change in progress. *Language and Speech*, 55(3), 1–23.
- Kraljic, T., Brennan, S. E., & Samuel, A. G. (2008). Accommodating variation: Dialects, idiolects, and speech processing. *Cognition*, 107(1), 54–81.
- Labov, W. (1994). *Principles of language change: Internal factors* (Vol. 1). Oxford, UK: Blackwell.
- Labov, W., Karen, M., & Miller, C. (1991). Near-mergers and the suspension of phonemic contrast. *Language Variation and Change*, 3(1), 33–74.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74(6), 431–461.
- Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54(5), 358–368.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), 1–36.
- Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, 4(5), 187–196.
- Lindblom, B., Guion, S., Hura, S., Moon, S.-J., & Willerman, R. (1995). Is sound change adaptive? *Rivista di Linguistica*, 7, 5–36.
- Mees, I., & Collins, B. (1982). A phonetic description of the consonant system of standard Dutch. *Journal of the International Phonetic Association*, 12, 2–12.
- Mitterer, H. (2009). *Research stuff*. Retrieved from <http://www.holgermitterer.eu/research.html>.
- Mitterer, H., & Ernestus, M. (2008). The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition*, 109(1), 168–173.
- Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R² and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface* 14: 20170213.
- Nakai, S. (1998). The effect of vowel prototypicality and extremity on discrimination sensitivity. *The Journal of the Acoustical Society of America*, 103, 2041–2042.
- Newman, R. (1997). *Individual differences and the link between speech perception and speech production*. PhD Dissertation, State University of New York.
- Newman, R. (2003). Using links between speech perception and speech production to evaluate different acoustic metrics: A preliminary report. *The Journal of the Acoustical Society of America*, 113(5), 2850–2860.
- Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5(1), 42–46.
- Ohala, J. J. (1981). The listener as a source of sound change. In C. S. Masek, R. A. Hendrick, & M. F. Miller (Eds.), *Papers from the Parasession on Language and Behavior* (pp. 178–203). Chicago, IL: Chicago Linguistic Society.
- Paliwal, K., Lindsay, D., & Ainsworth, W. (1983). Correlation between production and perception of English vowels. *Journal of Phonetics*, 11(1), 77–83.
- Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., & Zandipour, M. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *The Journal of the Acoustical Society of America*, 116(4), 2338–2344.
- Pierrehumbert, J. (2001). *Lenition and contrast* (Vol. 45). Amsterdam, Netherlands: John Benjamins Publishing.
- Pierrehumbert, J. (2002). Word-specific phonetics. *Laboratory Phonology*, 7, 101–139.
- Pierrehumbert, J. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, 46(2–3), 115–154.
- Pinget, A. (2015). *The actuation of sound change*. PhD Dissertation, Utrecht University, The Netherlands.

- Pinget, A., Rotteveel, M., & Van de Velde, H. (2014). Herkenning en evaluatie van regionaal gekleurd Standaardnederlands in Nederland [Identification and evaluation of regionally marked standard Dutch in the Netherlands]. *Nederlandse Taalkunde*, 19(1), 3–45.
- Pinget, A., Kager, R., & Van de Velde, H. (2016). Regional differences in the perception of a consonant change in progress. *Journal of Linguistic Geography*, 4(2), 65–75.
- Pisoni, D. (1997). Some thoughts on ‘normalization’ in speech perception. In K. Johnson & J. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9–32). New York, NY: Academic Press.
- Pisoni, D. (1975). Auditory short-term memory and vowel perception. *Memory & Cognition*, 3(1), 7–18.
- Repp, B. H. (1981). Two strategies in fricative discrimination. *Perception & Psychophysics*, 30(3), 217–227.
- Repp, B. H. (1984). Categorical perception: Issues, methods, findings. *Speech and Language: Advances in Basic Research and Practice*, 10, 243–335.
- Rietveld, A., & van Heuven, V. (2009). *Algemene fonetiek* [General phonetics]. Bussum, Netherlands: Coutinho.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192.
- Rojczyk, A. (2011). Perception of English voice onset time continuum by Polish learners. In J. Arabski & A. Wojtaszek (Eds.), *The acquisition of L2 phonology* (pp. 37–58). Bristol, Buffalo, Toronto: Multilingual Matters.
- Slis, I., & Cohen, A. (1969). On the complex regulating the voiced-voiceless distinction I and II. *Language and Speech*, 12, 80–102.
- Slis, I., & van Heugten, M. (1989). Voiced-voiceless distinction in Dutch fricatives. In H. Bennis & A. van Kemenade (Eds.), *Linguistics in the Netherlands* (pp. 123–132). Dordrecht: AVT Publications.
- Smakman, D. (2006). *Standard Dutch in the Netherlands: A sociolinguistic and phonetic description*. PhD Dissertation, Radboud University Nijmegen, Utrecht, The Netherlands.
- Smits, R., & van Alphen, P. (2004). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics*, 32(4), 455–491.
- Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language*, 60(4), 487–501.
- Taeldeman, J., & Hinskens, F. (2013). The classification of the dialects of Dutch. In F. Hinskens & J. Taeldeman (Eds.), *Language and space: An international handbook of linguistic variation, Volume 3 Dutch* (pp. 129–141). Berlin: Mouton de Gruyter.
- Thomas, E. R. (2000). Spectral differences in /ai/ offsets conditioned by voicing of the following consonant. *Journal of Phonetics*, 28(1), 1–25.
- Van Alphen, P., & Smits, R. (2004). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics*, 32(4), 455–491.
- Vandekerckhove, R. (2009). Dialect loss and dialect vitality in Flanders. *International Journal of the Sociology of Language*, 196–197, 73–97.
- van den Berg, R. (1988). *The perception of voicing in Dutch two-obstruent sequences* (PhD Dissertation). University of Nijmegen, The Netherlands.
- Van de Velde, H. (1996). *Variatie en verandering in het gesproken Standaardnederlands (1935–1993)* [Variation and change in spoken standard Dutch]. PhD Dissertation, University of Nijmegen, The Netherlands.
- Van de Velde, H., Gerritsen, M., & van Hout, R. (1996). The devoicing of fricatives in standard Dutch: A real-time study based on radio recordings. *Language Variation and Change*, 8(2), 149–175.
- Van de Velde, H., & Houtermans, M. (1999). Vlamingen en Nederlanders over de uitspraak van nieuwslezers [Flemish and Dutch evaluations of the pronunciation of news readers]. In H. Broekhuis & P. Fikkert (1999), *Artikelen van de Derde Sociolinguïstische Conferentie* (pp. 451–462). Delft, Netherlands: Uitgeverij Eburon.
- Van de Velde, H., & van Hout, R. (2002). Loan words as markers of differentiation. In H. Broekhuis & P. Fikkert (Eds.), *Linguistics in the Netherlands 2002* (pp. 163–173). Amsterdam, Netherlands: John Benjamins.
- Van de Velde, H., Kissine, M., Tops, E., van der Harst, S., & van Hout, R. (2010). Will Dutch become Flemish? Autonomous developments in Belgian Dutch. *Multilingua*, 29, 385–416.

- van der Harst, S. (2011). *The vowel space paradox. A sociophonetic study on Dutch*. PhD Dissertation, Radboud University Nijmegen, The Netherlands.
- van der Harst, S., Van de Velde, H., & van Hout, R. (2014). Variation in Standard Dutch vowels: The impact of formant measurements methods on identifying the speaker's regional origin. *Language Variation and Change*, 26(2), 247–272.
- van der Wal, M., & van Bree, C. (1992). *Geschiedenis van het Nederlands* [History of the Netherlands] (pp. 1–494). Utrecht: Het Spectrum.
- Van de Velde, H. (1996). *Variatie en verandering in het gesproken Standaardnederlands (1935–1993)* [Variation and change in spoken standard Dutch]. PhD Dissertation, University of Nijmegen, The Netherlands.
- Van de Velde, H., Gerritsen, M., & van Hout, R. (1996). The devoicing of fricatives in standard Dutch: A real-time study based on radio recordings. *Language Variation and Change*, 8(2), 149–175.
- van Son, R. (2000). Protocol voor het oplijnen van fonetische transcripties met spraak. Retrieved from <http://www.fon.hum.uva.nl/IFA-SpokenLanguageCorpora/IFAcopus/SLcorpus/LabelProtocol/LabelProtocol.pdf>.
- van Son, R., & Pols, L. (1996). An acoustic profile of consonant reduction. In *Proceedings of the Fourth International Conference on Spoken Language 3*, pp. 1529–1532.
- Verhoeven, J., & Hageman, G. (2007). De verstemlozing van fricatieven in Vlaanderen. *Nederlandse Taalkunde*, 12, 139–151.
- Ziliak, Z., & Van de Velde, H. (2008). Stop variation in Dutch. *Presentation at New Ways of Analyzing Variation*, 37, Houston, TX.
- Zwaardemaker, H., & Eijkman, L. (1928). *Leerboek der Phonetiek* [Handbook of phonetics]. Haarlem: De Erven F. Bohn.