

Regional differences in the perception of a consonant change in progress

Anne-France Pinget*, René Kager* and Hans Van de Velde[^]

* *Utrecht Institute of Linguistics OTS, Utrecht University*

[^] *Fryske Akademy, Leeuwarden*

This study aims at testing whether there are regional differences in the perception of the labiodental fricative contrast in Dutch. Previous production studies have shown that the devoicing of initial labiodental fricatives is a change in progress in the Dutch language area. We present the results of a speeded identification task in which fricative stimuli were systematically varied for two phonetic cues, voicing and duration. Listeners ($n = 100$) were regionally stratified, and the regions ($k = 5$) reflect different stages of this sound change in progress. Voicing turned out to be the strongest categorization cue in all regions; duration only played a minor role. Regional differences showed up in the perception of the consonantal contrast that matched regional differences in production reported in previous studies. The addition of random slopes in the mixed model regression showed the importance of within-regional variation.

Keywords: perception, sound change, fricatives, devoicing, regional variation, link perception-production

1. Introduction

Human speech perception is attuned to the listener's native language. Babies start organizing their perceptual space on the basis of the distributions of sounds present in their surroundings (e.g. Kuhl et al., 1992). As a result, adult listeners perceive speech sounds categorically. Early accounts argued that listeners try to relate a perceived sound to the phonological categories in their perceptual space (e.g., Liberman et al., 1957). Kuhl (1991) proposed an alternative model, the native language magnet theory, in which the phonetic categories of one's native language are organized in terms of prototypes of specific speech sounds. These prototypes function as "perceptual magnets" since they attract or assimilate phonetically similar members, and thus facilitate the processing of the variability in the speech input. More recently, this magnet effect has been shown to emerge from the exemplar-based organization of the perceptual space (Johnson, 1997; Goldinger, 1997) and to be a result of optimally solving the statistical problem of perception (Guenther & Gjaja, 1996; Feldman, Griffiths & Morgan 2009; Boersma & Hamann, 2008). Regardless of the theoretical account, it is clear that the language input plays a major role in the organization of our perceptual space. The consequence of this perceptual attunement to the input is that listeners with different native languages show different perceptual patterns, as they have been exposed throughout their life to different sound systems. This has been confirmed by numerous studies on both the

segmental (e.g., Escudero, Simon & Mitterer, 2012; Flege, Takagi & Mann, 1996; Guion et al., 2000) and the supra-segmental level (e.g., Guion, Harada & Clark, 2004; Hallé, Chang & Best, 2004; Hirata, 2004; McAllister, Flege & Piske, 2002).

An interesting question is whether similar perceptual differences show up in speakers of closely related language varieties, where the phonological differences are much smaller than between the language pairs studied until now. This can be tested with speakers of dialects or accents of the same language. Ladefoged and Broadbent (1957) showed that the identification of vowels differs between Scottish and English speakers of English, Willis (1972) between American and Canadian speakers of English. Thomas (2011:27) observed that the perceptual differences in Willis (1972) closely match the differences between Buffalo and Ontario accents. Janson (1983) found that speakers from different regions and age groups hear Swedish vowels differently. More recently, Kendall & Fridland (2012) studied regional differences in the perception of mid front tense and lax vowels in American English. Their results show that both regional affiliation and individual participation in regional shifts in production play a role in speech perception. Until now, studies have largely focused on vowels, and it is unclear whether the same kind of variation can also be found in consonantal contrasts. The perception of consonants has been shown to be more categorical than the perception of vowels (e.g., Fry et al., 1962; Pisoni, 1975; Repp, 1984), but we still need to deepen our understanding of their variation patterns in perception.

Moreover, synchronic regional patterns of variation commonly reflect diachronic patterns of change. In previous studies (with the exception of Kendall &

Corresponding author: Anne-France Pinget, a.c.h.Pinget@uu.nl, Tel: +31 30 253 6049, Utrecht Institute of Linguistics OTS, Universiteit Utrecht, Trans 10, 3512 JK Utrecht, The Netherlands.

Fridland, 2012), regional differences were assumed to be stable and were not described in the context of sound change. In the frame of research on diachronic change, however, it is crucial to determine whether this variation in speech perception plays a role in sound change.

In this study we focus on a well-researched consonantal change in progress in the Dutch language area: the devoicing of the initial labiodental fricative /v/. This paper presents the results of a speeded identification task aimed at getting insight in the speech perception patterns of speakers from five regions of the Dutch language area, which differ in their realization of labiodental fricatives and represent different stages of sound change. We expect that regional differences in the implementation of the /v/-/f/ contrast in production will be reflected in speech perception. Our study aims to answer the following research questions:

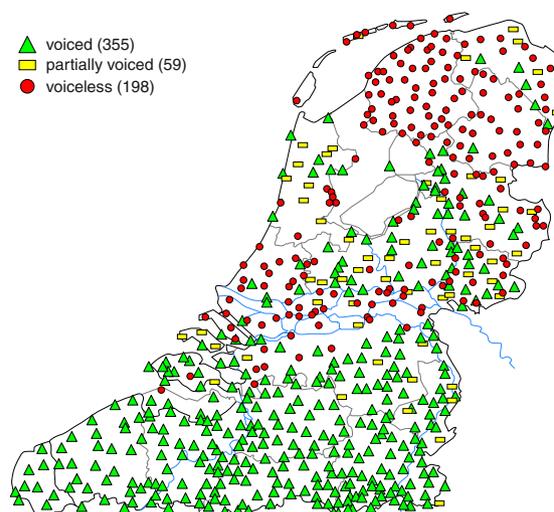
1. Does regional variation in the perception of labiodental fricatives mirror the previously observed regional variation in production in the Dutch language area?
2. To what extent are individual perceptual patterns homogeneous within a region? What is the nature of within-region variation in the perception of labiodental fricatives?

2. Insights from production studies

Dutch /v/ derives from West-Germanic /f/, which became voiced in the oldest phase of Dutch in non-final positions (Gussenhoven & Bremmer, 1983) and merged with /v/ in word internal onset position that derived from a West-Germanic voiced bilabial fricative. In present-day standard Dutch the labiodental fricatives /v/ and /f/ are contrastive in onset position. In coda position, voiced /v/ does not occur as a result of final devoicing. During the last decennia, it was frequently observed that the voiced-voiceless distinction is neutralized by many speakers of standard Dutch, resulting in the use of the voiceless variant. The phenomenon is said to be a merger or near-merger, in which the sound change leads to the collapse of the phonemic contrast, so that the two previously distinct phonemes merge into a single phoneme (see Labov, 1994). This devoicing is a well-researched phenomenon from the point of view of phonetics (Debrock, 1977, 1978; Slis & Cohen, 1969; Van den Berg & Slis, 1985; Van den Berg, 1989), dialectology (Van Reenen & Wattel, 1992), and language change in progress (Van de Velde, Gerritsen & van Hout, 1996; Kissine, Van de Velde & van Hout, 2003). The functional load of labiodental fricatives is very low. Dutch only has a few minimal pairs with initial /v/ and /f/ (approximately ten) and most of these have a very low frequency.

The major phonetic cue for the voiced/voiceless distinction in the production of Dutch fricatives is the presence or absence of vocal cord vibration (Slis & Cohen, 1969; Van den Berg, 1998). In addition, Dutch voiceless fricatives tend to be longer than their voiced counterparts (Slis & Cohen, 1969; Slis & Van Heugten, 1989).

Production studies showed that—in standard Dutch—/v/ is almost completely devoiced (and merged with /f/) in the North of the Netherlands, strongly devoiced in the central area of the Netherlands, moderately devoiced in the South of the Netherlands, and incipiently devoiced in Flanders (Van de Velde, Gerritsen & van Hout, 1996; Kissine, Van de Velde & van Hout, 2003). In addition to the regional differences in the amount of voicing of /v/, differences in the implementation of the duration contrast between /v/ and /f/ were observed (Kissine et al., 2003). The differences in the standard language can be linked to older dialectal patterns. This can be illustrated with a map of the pronunciation of word initial /v/ in the Dutch and Frisian dialects, based on the item *verf* ('paint') in the Goeman-Taeldeman-Van Reenen dataset (<http://www.meertens.knaw.nl/mand/database/>). After recoding the different transcriptions to three variants (voiced, partially voiced, and voiceless) Map 1 was generated by means of the Meertens Mapper (<http://www.meertens.knaw.nl/projecten/mand/CARTkartografieapp.html>). A north-south pattern becomes obvious: in the south (i.e., Flanders and the southern area in the Netherlands), voiced realizations show up; voiceless realizations are found in the north. In some dialects of the northern part, words starting with voiced fricatives were originally realized with voiceless fricatives in Middle Dutch (Gussenhoven & Bremmer, 1983; van Reenen &



Map 1. Pronunciation of /v/ in *verf* in Dutch and Frisian dialects.

Wattel, 1992). However, it is unclear how this influenced their use in standard Dutch.

Partially voiced realizations are mainly found in a transition zone in the central area of the Netherlands, next to voiced and voiceless realizations. In the dialects spoken in Flanders, only voiced realizations of /v/ were attested. The recordings of the Goeman-Taeldeman-Van Reenen project were made in the early 1990s. At the same time, the first signs of devoicing are attested in Belgian standard Dutch spoken by broadcasters (Van de Velde et al., 1996). Ten years later devoicing of /v/ is observed more widely in the speech of Dutch language teachers (Kissine et al., 2003).

On the basis of the above-mentioned production studies, five regions of the Dutch language area were chosen to represent different stages of the devoicing of /v/:

- Groningen (GR): in the North of the Netherlands, where the sound change is almost complete, with a (near)-merger of /v/ and /f/. Almost all /v/ realizations are aperiodic and there are no duration differences between /v/ and /f/;
- South-Holland (SH): in the West of the Netherlands, where the sound change is in an advanced stage, with strong devoicing of /v/. Fully periodic /v/ realizations are rare, /v/ realizations are shorter than /f/;
- Limburg (LI): in the South of the Netherlands, where the sound change is in an intermediate stage, with weak devoicing of /v/. Fully periodic /v/ realizations are still present, /v/ realizations are shorter than /f/;
- Flemish-Brabant (BR): in the centre of Belgium, where the sound change is in an incipient phase and still unnoticed by language users, with an incipient devoicing of /v/, and /v/ realizations being much shorter than /f/.
- West-Flanders (WF): in the North-West of Belgium, where the sound change is in an incipient phase and still unnoticed by language users, with only small amounts of devoicing of /v/, and /v/ realizations being much shorter than /f/.

Although these five regions are considered in our study to represent steps in the progress of the sound change, one should be aware of the fact that Flemish-Brabant and West-Flanders are in Flanders, which has been shown to develop its own variety of the standard norm, Belgian Standard Dutch, whereas Groningen, South-Holland, and Limburg are in the Netherlands, where Netherlandic Standard Dutch is spoken (Van de Velde et al., 1997).

Besides these regional differences, individual perceptual differences might be expected. In production studies, it is often the case that some individuals show progressive patterns, largely adopting the new variant and therefore leading the change, while others show conservative patterns, being apparently reluctant to the

change in progress. Individual differences between and within regions typically form a continuum of variation, with overlapping speakers between adjacent regions. The second research question aims to investigate whether this is also the case in perception.

3. Method

3.1 Listeners

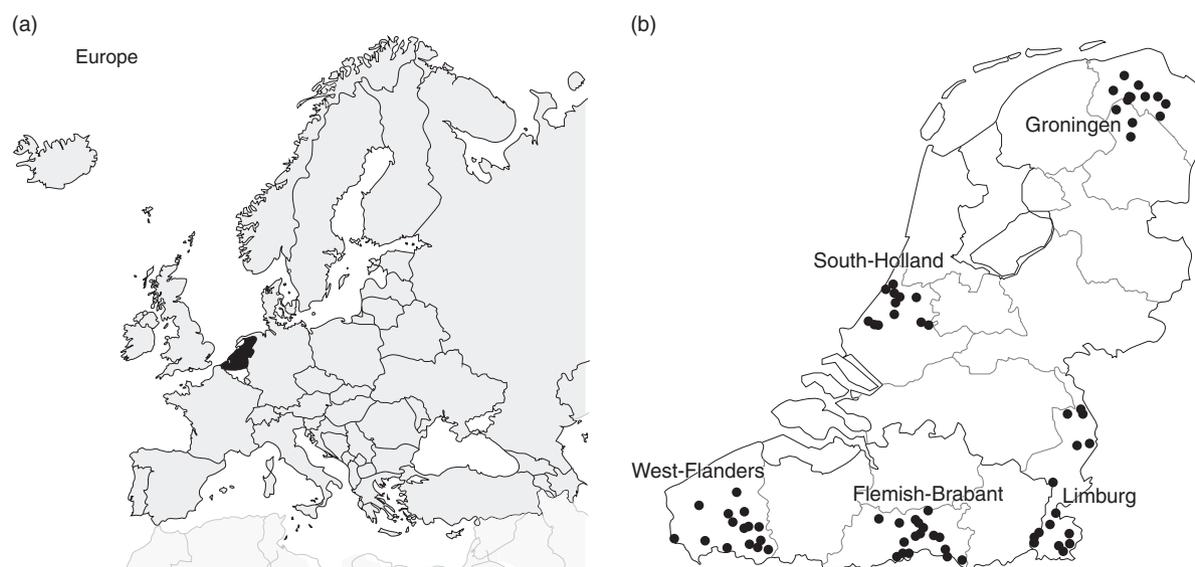
For each region described in Section 2, twenty native speakers of Dutch, born and raised in this region, were selected. The participants were equally stratified for gender (10 males and 10 females). The factors age and educational level were kept constant: all participants were highly educated young adults (they were attending or recently graduated from university or non-university higher education), aged between 18 and 28 years. In Map 2, the hometown of each participant is presented.

3.2 Stimuli

Based on the sociophonetic insights from studies reviewed in the introduction, we chose two phonetic dimensions along which we investigated the perception of the devoicing of /v/: VOICING (the degree of voicing) and DURATION. A two-dimensional speech continuum between /v/ and /f/ was generated by manipulating these two phonetic dimensions. Fricatives were presented in a CV syllable. As the /i/ vowel shows least regional variation in vowel quality in Standard Dutch (Van der Harst, 2011: 159), we opted for this vowel to avoid a bias caused by regional differences in the perception of the vowel.

For the creation of the /vi/-/fi/ continuum, naturally produced syllables /vi/ and /fi/ were used. For ecological validity, we choose to use and manipulate natural, spoken stimuli instead of synthesized ones. Speech material of a male native speaker of Dutch (25 years old, from the South-Holland region), who is a trained phonetician, was digitally recorded with a sample frequency of 44.1 kHz in a sound-attenuated cabin. The speaker pronounced a series of one-syllable Dutch non-words, which included the target combinations /vi/ and /fi/. The duration of the selected fully voiced (periodic) [v] was 184 ms; the fully voiceless (aperiodic) [f] 176 ms. The vowel [i] had a duration of 251 ms in the /vi/ realization and 239 ms in the /fi/ realization. The mean F0 value of [v] was 131 Hz.

The first dimension—voicing—is the percentage of pitch present in the fricatives. It ranges from 100%, a fully voiced fricative containing pitch over the whole segment, to 0%, a fully unvoiced fricative with no trace of pitch. Along this dimension, the continuum consists of nine steps: 100%, 87.5%, 75%, 62.5%, 50%, 37.5%, 25%, 12.5% and 0% voicing. The second dimension is



Map 2. Map of (a) the Netherlands within Europe and (b) the five selected regions in the Dutch language area: West-Flanders, Flemish-Brabant, Limburg, South-Holland and Groningen. Each point represents the home town of one or more participants.

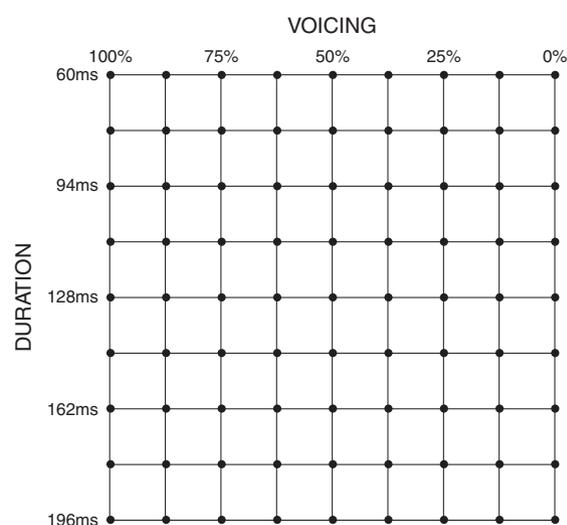


Figure 1. Schematic representation of the two bi-dimensional fricative continua, varying voicing horizontally and duration vertically.

the duration of the fricative. It ranges from 60 to 196 ms, with an equal interval of 17 ms (nine steps). This resulted in a two-dimensional grid consisting of nine steps for each dimension, thus a /vi-/fi/ continuum with 81 realizations (see Figure 1).

In order to generate the above-described continuum, the fricatives of the source recordings were extracted from their original context, lengthened to a duration of 196 ms, 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. In this way, we obtained the first nine sound realizations of the continuum with a constant duration of 196 ms and varying degrees of voicing (ranging from 0% to 100% of voicing). The remaining realizations were obtained by manipulating the duration of the fricatives. The syllable stimuli were obtained by concatenating the /v-/f/ realizations with the [i] produced in the original /vi/ context. To keep the vowel as constant as possible, it was shortened to 110 ms and the pitch contour of [i] was manipulated through the PSOLA pitch manipulation and LPC resynthesis functions of Praat. In the vowel transition, the pitch contour was flattened to values ranging between 130 and 135 Hz. From 60 ms after the beginning after the vowel onset, the pitch contour was gradually reduced to 120 Hz for the remaining duration of the vowel in order to get a falling contour.

3.3 Procedure

Participants listened in a sound-attenuated booth to the stimuli through a headphone Beyerdynamic DT 250. They were asked to categorize the realizations as being an /f/ or a /v/ sound by pressing the red or blue button of a button box labelled with the corresponding sound. The order of presentation of the consonants on the button box (i.e., f-v or v-f) was balanced between participants. The task was auto-paced. Reaction times (RT) were recorded from the beginning of the stimuli. A participant disposed of a time window of 800 ms after

the end of the stimulus to give their response. A response given after this time window was not recorded to avoid responses resulting from second guesses, and the participant automatically got a new stimulus. Twelve stimuli were presented in the practice session to familiarize the participant with the task. After the practice session, the test phase started. The 81 stimuli were randomly presented five times.

The binary responses obtained in this speeded forced-choice identification task were analysed with logistic regressions with the percentage of identification as /v/ as the dependent variable. As commonly used in the analysis of categorical perception (e.g., Kendall & Fridland, 2012), logistic regressions provide psychometric curves that are defined by 1) their slopes, which indicate how categorical the judgment is (the steeper the curve, the more categorical the judgment) and 2) their medians, which represent the categorical boundary between the /v/ and /f/.

3.4. Hypotheses

Considering the regional differences in production sketched above, we formulate the following hypotheses related to variation between regions.

- (1) Voicing will be a stronger cue for the identification of fricatives than duration.
- (2) The less devoicing of /v/ in production per region, the more categorically the stimuli will be perceived. We predict the following order for the steepness of the slope for both voicing and duration: WF > BR > LI > SH > GR.
- (3) The less devoicing of /v/ in production per region, the closer the categorical boundary between categories will be to the right of the /v/-/f/ continuum. The expected order is WF > BR > LI > SH > GR.

We formulate the following hypotheses related to perceptual variation within regions.

- (4) Variation within regions will be the largest around the categorical boundary and the smallest at the extremes of the continuum.
- (5) Variation will be small in regions where sound change is almost completed and in regions where sound change is still incipient. In contrast, variation will be large in regions where sound change is in an intermediate stage.

4. Results

Trials with response times shorter than the duration of the consonants were excluded from the analysis, as they occurred before the participants heard the entire first segment. Hearing the entire first segment was crucial,

Table 1. Centralization of the continuum along both dimensions.

Continuum steps	1	2	3	4	5	6	7	8	9
Centralized numeric variable	-4	-3	-2	-1	0	1	2	3	4
Voicing dimension	fully voiced				fully voiceless				
Duration dimension	short					long			

since duration was a manipulated cue in the fricatives. Moreover, responses outside the time window (800 ms after the end of a stimulus) were treated as missing values. Excluded data (too short latency) and missing values (too long latency) represent 6.97% of the data. In total, 37,678 valid observations were used for the quantitative analysis.

In order to obtain parameters that are more interpretable within logistic regression, the continuum steps along both dimensions were centralized, so that step 5 would be equal to 0. The centralized continuum thus ranged from -4 to 4 instead of from step 1 to step 9 (see Table 1). Along the voicing dimension, the leftmost part (negative values) refers to the most periodic realizations and the rightmost part (positive values) refers to the less periodic realizations. Along the duration dimension, the negative values refer to the shortest realizations and the positive values refer to the longest realizations.

4.1. Regional differences

A mixed-effect logistic regression was fitted to the data. The models included listeners ($n = 100$) as random intercept and regions and continuum steps (both voicing or duration) as fixed factors. Voicing and duration were added as a numeric variable. In logistic regression, the probability of x ($P(x)$)—in this case the probability of a /v/ response—is predicted by the following equation:

$$P(x) = \frac{e(\beta_0 + \beta_1 x)}{1 + e(\beta_0 + \beta_1 x)}$$

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. The median (i.e. the point where the probability $P(x)$ is equal to 0.5) can be calculated on the basis of these two estimates with the following formula:

$$\text{Median} = \frac{-\beta_0}{\beta_1}$$

The regional differences in perceptual patterns are shown in Figure 2. The estimates of the model and their significance are presented in Table 2.

The mixed model logistic regression showed that there is a significant effect of voicing. This slope is

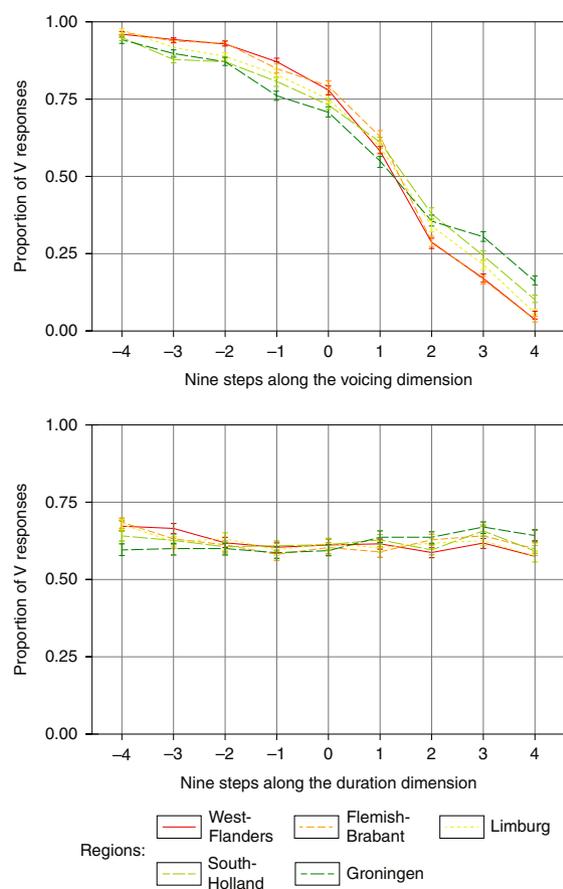


Figure 2. Results of the identification task split up by region. The centralized nine-steps continuum along the voicing dimension (in the upper panel) and along the duration dimension (in the lower panel) are presented on the x-axis and the proportion of /v/ responses on the y-axis. Error bars represent ± 1 standard error.

negative (-1.040) since the more to the left (the more voicing), the greater the /v/ responses. Furthermore, regions significantly differ from each other. Differences in slopes between regions are shown by the significant interactions between voicing and regions. All regions significantly differ from West-Flanders, which is taken as the reference level in the intercept, since this region was expected to show the most conservative perceptual patterns. By changing the reference level systematically in the analyses, it was shown that all regions differ significantly from each other. West-Flanders has the steepest slope along the voicing dimension (-1.117), followed by Flemish-Brabant (-0.928). Limburg (-0.725) and South-Holland (-0.613) have a less steep slope and Groningen has the gentlest slope of all regions (-0.546). West-Flemish listeners almost categorically respond /v/ in the first three steps of the continuum (-4 to -2), while the other regions start lower and gradually decrease along the continuum. The 0.5 cut-off point is situated to the right of our continuum (between

centralized step +1 and +2), but does not significantly differ across regions, which is surprising as there are—in production—large differences in the degree of voicing of /v/ between the regions (Kissine et al. 2003).

Furthermore, we observe that there are no 0.5 cut-off points for duration. For all regions, the slopes are almost flat and the model shows that there is no significant effect of duration. Participants do not seem to use duration as a cue when categorizing the sounds. However, Flemish-Brabant and Groningen significantly differ from West-Flanders (the reference level) as they show significant interactions. Only Groningen (when taken as reference level) shows a slope that is significantly higher than 0. This slightly significant positive slope in Groningen means that these listeners used the duration cue in the opposite way as expected: the longer the duration of the sound, the higher the chance to categorize it as /v/. Listeners from other regions did not make use of the duration cue.

Finally, we observe a significant but weak interaction between voicing and duration. The conjunctions of both dimensions strengthen the regional effect that we found for voicing. Except for Flemish-Brabant, all regions significantly differ from West-Flanders in decreasing order of the steepness of the slope: Limburg, South-Holland and Groningen.

4.2. Individual differences within regions

In this section, individual differences in the speech perception of labiodental fricatives within regions are investigated. In the previous section, the statistical analysis revealed consistent differences between regions. Regions significantly differed from each other in their use of voicing. In every region, however, there appeared to be large individual differences (see Figure 3). This poses the question of the customization of the statistical analysis to the research questions. Barr et al. (2013) for instance argue in favour of the addition of random slopes, next to the random intercept in a mixed model: “linear mixed-effects models generalize best when they include the maximal random effects structure justified by the design” (Barr et al., 2013: 255). The model presented in the previous section only contained random intercepts. However, when adding random slopes (i.e., allowing the perceptual slope to vary between participants within the same region), some of the significant differences between regions disappeared: the three Dutch regions (South-Holland, Limburg, and Groningen) still differ from the Belgian ones (West-Flanders and Flemish-Brabant), but the Dutch regions do not differ from each other, nor do the Belgian regions (see Table 3), indicating that listeners show a lot of variation within regions and overlap between regions. Although devoicing of /v/

Table 2. Results of the mixed-effects logistic regressions with voicing, duration, and region as fixed factors and listeners as random intercept. The region of West-Flanders is taken as the reference level. * The asterisk indicates significance ($\alpha = 0.05$).

Random effects:	Variance	SD		
Participant (Intercept)	0.443	0.666		
Fixed effects:	Estimate	Std. Error	z value	p value
Voicing	-1.040	0.024	-43.920	<0.001*
Duration	-0.013	0.216	-0.780	0.435
West-Flanders	1.262	0.155	8.140	<0.001*
Flemish-Brabant	1.194	0.155	7.730	<0.001*
Limburg	0.852	0.153	5.570	<0.001*
South-Holland	0.723	0.152	4.750	<0.001*
Groningen	0.635	0.152	4.170	<0.001*
Voicing * Flemish-Brabant	0.110	0.032	3.460	<0.001*
Voicing * Limburg	0.277	0.029	9.510	<0.001*
Voicing * South-Holland	0.385	0.028	13.710	<0.001*
Voicing * Groningen	0.445	0.028	16.010	<0.001*
Duration * Flemish-Brabant	0.051	0.022	2.290	0.022*
Duration * Limburg	-0.011	0.021	-0.550	0.583
Duration * South-Holland	0.024	0.020	1.180	0.239
Duration * Groningen	0.088	0.020	4.350	<0.001*
Voicing * Duration	-0.077	0.008	-9.040	<0.001*
Voicing * Duration * Flemish-Brabant	0.006	0.011	0.510	0.610
Voicing * Duration * Limburg	0.038	0.010	3.640	<0.001*
Voicing * Duration * South-Holland	0.042	0.010	4.170	<0.001*
Voicing * Duration * Groningen	0.049	0.010	4.950	<0.001*

started showing up in Belgian Dutch (Van de Velde et al., 1996; Kissine et al., 2003), perceptual differences between speakers of the respective national varieties are clearly present.

We will now investigate these individual differences more deeply. To get more insight into the between-listener variation within regions, we look first at the standard deviation within each region on each step of the continuum (i.e., a measure of data dispersion within each region) in Table 4 and second, at a boxplot showing the proportion of /v/ responses split up by the steps of the voicing continuum and region in Figure 3. The median is represented by the horizontal line, the first and third quartiles by the white box, and the 95% confidence interval of the median by the vertical line. Outliers are marked by dots.

From both Table 4 and Figure 3, it is clear that between-listener variation is largest in step +1 and in step +2 of the continuum. As shown in Section 4.1, these steps match the perceptual boundary between /v/ and /f/. The standard deviations within each region for these steps swing around 0.5 as compared to 0.2–0.3 for steps at the extremes of the continuum. In steps -4 and -3, almost all participants give around 100% /v/ responses. In South-Holland and Groningen, the individual variation is larger. For steps -2, -1 and 0,

individual variation increases in all regions. In steps +2 and +3, all regions roughly show comparable individual variation. In step +4, the variation is very small, except for South-Holland and Groningen.

Moreover, it appears that the standard deviations of the Dutch regions are systematically larger than those of the Flemish regions. We hypothesized that individual variation would be small both in regions where sound change is almost completed and in regions where the sound change is still incipient. In both Table 4 and Figure 3, it turns out that the largest between-listener variation is to be found in the regions South-Holland and Groningen, and this on each step of the continuum.

5. Discussion and Conclusion

This paper presented the results of a speeded identification task performed by one hundred native speakers of Dutch from five different regions. The experiment aimed at getting insight in the regional differences in the identification of labiodental fricatives, a merger in progress in Dutch.

Our first goal was to figure out which phonetic cues the listeners rely on when categorizing labiodental fricatives. Based on speech production studies (Slis & Cohen, 1969; Slis & Van Heugten 1989; Van den Berg,

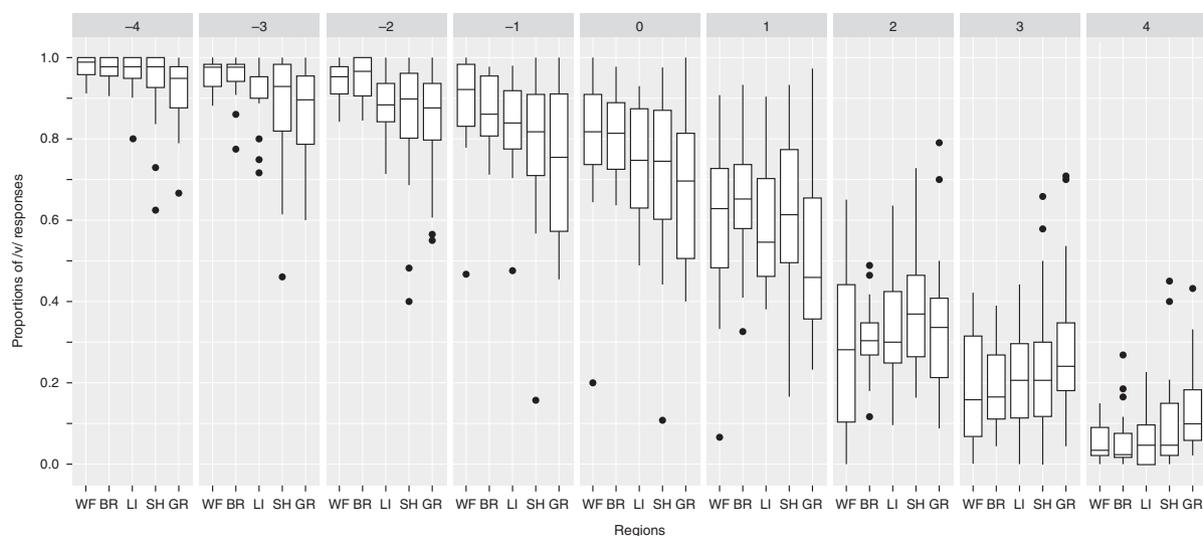


Figure 3. Boxplot of the proportion of /v/ responses split up by the steps of the voicing continuum and by regions.

Table 3. Results of the mixed-effects logistic regressions with voicing, duration and region as fixed factors, listeners as random intercept, and voicing and duration by listeners as random slopes. The region of West-Flanders is taken as the reference level. * The asterisk indicates significance ($\alpha = 0.05$).

Random effects:	Variance	SD		
Participant (Intercept)	0.832	0.912		
Voicing by participant (Slope)	0.093	0.308		
Duration by participant (Slope)	0.016	0.127		
Fixed effects:	Estimate	Std. Error	z value	p value
Voicing	-1.124	0.073	-15.353	<0.001*
Duration	-0.031	0.033	-0.919	0.358
West-Flanders	1.380	0.210	6.563	<0.001*
Flemish-Brabant	1.321	0.209	6.318	<0.001*
Limburg	0.987	0.208	4.754	<0.001*
South-Holland	0.983	0.208	4.725	<0.001*
Groningen	0.848	0.208	4.067	<0.001*
Voicing * Flemish-Brabant	0.104	0.103	1.014	0.310
Voicing * Limburg	0.258	0.102	2.528	0.011*
Voicing * South-Holland	0.335	0.102	3.280	<0.001*
Voicing * Groningen	0.424	0.102	4.162	<0.001*
Duration * Flemish-Brabant	0.064	0.047	1.377	0.168
Duration * Limburg	0.008	0.046	0.171	0.864
Duration * South-Holland	0.047	0.046	1.032	0.301
Duration * Groningen	0.120	0.046	2.614	0.009*
Voicing * Duration	-0.077	0.009	-8.618	<0.001*
Voicing * Duration * Flemish-Brabant	0.004	0.012	0.318	0.751
Voicing * Duration * Limburg	0.034	0.011	3.027	0.002*
Voicing * Duration * South-Holland	0.037	0.011	3.362	<0.001*
Voicing * Duration * Groningen	0.050	0.011	4.729	<0.001*

1998; Kissine et al., 2003), two crucial dimensions in the implementation of the voiced/voiceless contrast in labiodental fricatives—voicing and duration—were chosen to test perception. Hypothesis 1 was confirmed, as all participants consistently used voicing as the main

cue in their identification of labiodental fricatives: the more voicing in the fricative, the higher the /v/ responses. Duration turned out to play a minor role in the perception of fricatives, as there was no significant effect of duration and only a weak interaction with voicing.

Table 4. Standard deviations calculated for each step of the continuum (ranging from -4 to 4) and within each region.

	-4	-3	-2	-1	0	1	2	3	4
West-Flanders	0.15	0.20	0.23	0.32	0.40	0.49	0.46	0.39	0.23
Flemish-Brabant	0.16	0.21	0.22	0.34	0.39	0.48	0.46	0.39	0.23
Limburg	0.17	0.27	0.32	0.37	0.43	0.49	0.47	0.41	0.24
South-Holland	0.23	0.34	0.34	0.40	0.45	0.49	0.48	0.42	0.30
Groningen	0.27	0.33	0.35	0.44	0.46	0.50	0.47	0.45	0.35

This absence of the duration effect might mean either that listeners effectively did not use duration as a phonetic cue to categorize fricatives, or that this result is an artefact of the type of experiment. In this kind of phonetic experiment, it is indeed possible that participants did not rely on the duration cues. Since the CV stimuli were presented in isolation, participants might not have been able to normalize for duration against other surrounding syllables, as it is normally done in spontaneous speech. Consequently, they might have shut down the use of the duration cue. As a result, it might not be entirely clear from this experiment what role duration plays in the perception of the voicing contrast in Dutch /v/ and /f/. However, Kissine et al. (2003) showed that speakers from the strongly devoicing South-Holland and Groningen regions did no longer make a durational difference between /v/ and /f/. Furthermore, the production data of the listeners in this study show a similar tendency: speakers from Groningen almost do not differ in duration between /v/ and /f/ (Pinget 2015:50). This shows that for speakers with strong devoicing of /v/, duration is no longer used to distinguish /v/ and /f/ in production. So, it should be no surprise that it does not show up in perception.

Second, the experiment aimed at revealing regional differences in the perception of the /v/-/f/ contrast. The five regions were chosen to represent different stages of the sound change in progress, ranging from Groningen, where there is almost complete devoicing of /v/, to West-Flanders, where the devoicing of /v/ is in an incipient phase. Listeners showed significant regional differences in the way they used the main cue, voicing. Hypothesis 2 was confirmed, since the perceptual patterns in West-Flanders were characterized by a steep psychometric curve (thus a more categorical perception), Groningen by a gentle slope (thus a less categorical contrast and more merged categories), and the other regions in between these two extremes. It should be noted that the perception slopes are not totally flattened—even in Groningen where devoicing is very advanced and has resulted in almost fully merged categories in production (Kissine et al., 2003). Listeners from strongly devoiced regions maintain traces of categorical perception. It seems that for these

listeners, the change in perception has not reached completion yet. It is very likely that these listeners maintain some perceptual contrast through speech input they receive from outside their region in which the contrast is still present. Dutch orthography, which has the /v/-/f/ contrast in word initial position, also contributes to the awareness of the contrast. Hypothesis 3, however, was not confirmed: we did not find regional differences in the categorical boundary between /v/ and /f/.

From the comparison between a statistical model with random intercepts only and a model with both random intercepts and random slopes, it became clear that individual differences in this type of research are important and should not be overlooked. For this reason, the second part of the analyses was devoted to variation within regions. It appeared that variation within regions was the largest around the categorical boundary and the smallest at the extremes of the continuum. Hence, Hypothesis 4 was confirmed. Individuals thus do not only differ in their categorization slope as turned out from the analysis above, but they also differ in how they categorize sounds around the boundary between the two categories.

From the analysis of standard deviations, it turned out that the largest individual variation was found in South-Holland and Groningen. This partly confirms Hypothesis 5: individual differences are relatively small in regions where sound change is still incipient. However, it refutes the hypothesis that individual variation is small in regions where sound change is almost completed. We hypothesize that this might be due to larger individual differences both in speech production and speech perception in these regions, and are planning to investigate this further by linking the production and perception data at the individual level.

In conclusion, we showed that there are regional differences in the speech perception of labiodental fricatives in the Dutch language. Perceptual variation in the context of sound change is thus not limited to vowels. Subtle but significant differences both between and within regions were reported. These differences are easily explainable within exemplar-based paradigms, since the configuration of the perceptual space is tuned

to the speech input received by a speaker/listener in his specific region. The traces left by the voiced exemplars are weaker and therefore more difficult to activate in the brain of individuals born and raised in a region where the devoicing is more advanced. The regional differences were mainly related to the perception slope, thus to the strength of the categoricalness of the contrast, and not to the categorical boundary. Exemplar-based paradigms would explain this fact by a weakening of the bimodal distribution, whereas—within the native language magnet theory—it would be interpreted as the gradual decrease in strength of the voiced prototype.

The differences between regions appeared to match production patterns reported in previous studies, pointing out a relationship between speech production and speech perception. Regional differences in the perception of a consonantal contrast undergoing sound change are present even in a highly standardized language like Dutch.

Acknowledgements

We gratefully acknowledge the anonymous reviewers who provided helpful comments and critiques on an earlier version of the manuscript. 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 also thank Hans Rutger Bosker, whose voice was used for the construction of the stimuli, Theo Veenker, for his technical assistance in recording and programming, and Mattis van den Bergh and Hugo Quené for their statistical advice.

Funding

This study is financed by the Netherlands Organisation for Scientific Research (NWO, 322-75-002).

References

- Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255-278.
- Benkí, José R. 2005. Perception of VOT and first formant onset by Spanish and English speakers. In James Cohen, Kara McAlister, Kellie Rolstad, Jeff MacSwan (eds.), *Proceedings of the 4th International Symposium on Bilingualism*, 240-248. Somerville, MA: Cascadilla Press.
- Boersma, Paul & Silke Hamann. 2008. The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25. 217-270.
- Boersma, Paul & David Weenink. 2014. *Praat: doing phonetics by computer*. <http://www.praat.org/> (2 December 2014).
- Debrock, M. 1977. An acoustic correlate of the force of articulation. *Journal of Phonetics* 5. 61-80.
- Debrock, M. 1978. Is the fortis-lenis feature really redundant in Dutch? *Leuvense Bijdragen* 67. 457-472.
- Escudero, Paola & Paul Boersma. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26. 4):551-585.
- Escudero, Paola, Ellen Simon & Holger Mitterer. 2012. The perception of English front vowels by North Holland and Flemish listeners: Acoustic similarity predicts and explains cross-linguistic and L2 perception. *Journal of Phonetics* 40. 280-288.
- Feldman, Naomi H., Tom L. Griffiths & James L. Morgan. 2009. The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological Review* 116. 4):752-782.
- Flege, James Emil, Naoyuki Takagi & Virginia Mann. 1996. Lexical familiarity and English-language experience affect Japanese adults' perception of /r/ and /l/. *Journal of the Acoustical Society of America* 99. 1161-1173.
- Fox, Robert Allen, James Emil Flege & Murray J. Munro. 1995. The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustical Society of America* 97(4). 2540-2550.
- Fry, Dennis Butler, Arthur S. Abramson, Peter D. Eimas & Alvin M. Liberman. 1962. The identification and discrimination of synthetic vowels. *Language and Speech* 5(4). 171-189.
- Gandour, Jackson T. & Richard A. Harshman. 1978. Crosslanguage differences in tone perception: A multidimensional scaling investigation. *Language and Speech* 21(1). 1-33.
- Goldinger, Stephen D. 1997. Words and voices—perception and production in an episodic lexicon. In Keith Johnson & John W. Mullennix (eds.), *Talker Variability in Speech Processing*, 33-66. San Diego: Academic Press.
- Guenther, Frank H. & Marin N. Gjaja. 1996. The perceptual magnet effect as an emergent property of neural map formation. *Journal of the Acoustical Society of America* 100. 1111-1121.
- Guion, Susan G., James Emil Flege, Reiko Akahane-Yamada & Jesica C. Pruitt. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America* 107. 2711-2724.
- Guion, Susan G., Tetsuo Harada & J. J. Clark. 2004. Early and late Spanish-English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition* 7. 207-226.
- Gussenhoven, Carlos & Rolf H. Bremmer. 1983. Voiced fricatives in Dutch: Sources and present-day usage. *North-Western European Language Evolution* 2. 55-71.
- Hallé, Pierre A., Yueh-Chin Chang & Catherine T. Best. 2004. Identification and discrimination of Mandarin Chinese tones

- by Mandarin Chinese vs. French listeners. *Journal of Phonetics* 32. 395-421.
- Hirata, Yukari. 2004. Training native English speakers to perceive Japanese length contrasts in word versus sentence contexts. *Journal of Acoustical Society of America* 116(4). 2384-2394.
- Janson, Tore. 1983. Sound change in perception and production. *Language*. 18-34.
- Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In Keith Johnson & John W. Mullennix (eds.), *Talker Variability in Speech Processing*. 145-165. San Diego: Academic Press.
- Kendall, Tyler & Valerie Fridland. 2012. Variation in the production and perception of mid front vowels in the US Southern Vowel Shift. *Journal of Phonetics* 40(2). 289-306.
- Kissine, Mikhail, Hans Van de Velde & Roeland van Hout. 2003. The devoicing of fricatives in standard Dutch. In Paula Fikkert & Leonie Cornips (eds.), *Linguistics in the Netherlands*, 2003. 93-104. Amsterdam: John Benjamins.
- Kuhl, Patricia K. 1991. Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics* 50(2). 93-107.
- Kuhl, P. K., Karen A. Williams, Francisco Lacerda, Kenneth N. Stevens & Björn Lindblom. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255. 606-608.
- Labov, William. 1994. *Principles of language change: Internal factors* (Vol. 1). Oxford: Wiley-Blackwell.
- Ladefoged, Peter & Donald E. Broadbent. 1957. Information conveyed by vowels. *Journal of the Acoustical Society of America* 29(1). 98-104.
- Lieberman, Alvin M., Katherine S. Harris, H. S. Hoffman & B. C. Griffith. 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54(5). 358.
- McAllister, Robert, James Emil Flege & Thorsten Piske. 2002. The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics* 30(2). 229-258.
- Mitterer, Holger. 2009. Research stuff. <http://www.holgermitterer.eu/research.html>. (2 December 2014).
- Pinget, Anne-France. 2015. *The actuation of sound change*. Utrecht: Utrecht University, LOT series PhD Thesis.
- Pisoni, David B. 1975. Auditory short-term memory and vowel perception. *Memory & Cognition* 3(1). 7-18.
- Repp, Bruno H. 1984. Categorical perception: Issues, methods, findings. *Speech and Language: Advances in basic research and practice* 10. 243-335.
- Slis, Iman H. & Antonie Cohen. 1969. On complex regulating voiced-voiceless distinction. *Language and Speech* 12(2). 80-102.
- Slis, Iman H. & Marieke van Heugten. 1989. Voiced-voiceless distinction in Dutch fricatives. *Linguistics in the Netherlands* 6. 123-132.
- Thomas, Erik R. 2011. Collecting data on phonology. In W. Maguire & A. McMahon (eds.), *Analysing Variation in English*, 7-29. Cambridge: Cambridge University Press.
- Van den Berg, R. J. H. 1989. Perception of voicing in Dutch two-obstruent sequences: Covariation of voicing cues. *Speech Communication* 8. 17-25.
- Van den Berg, R. J. H. & Iman H. Slis. 1985. Perception of Assimilation of Voice as a Function of Segmental Duration and Linguistic Context. *Phonetica* 42. 25-38.
- Van der Harst, Sander. 2011. *The vowel space paradox. A sociophonetic study on Dutch*. Utrecht: Utrecht University, LOT series PhD Thesis.
- Van de Velde, Hans, Marinel Gerritsen & Roeland van Hout. 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, Hans, Roeland van Hout & Marinel Gerritsen. 1997. Watching Dutch Change. *Journal of Sociolinguistics* 1(3). 361-391.
- Van Reenen, Pieter T. 1994. Driemaal /r/ in de Nederlandse dialecten. *Taal en Tongval* 46. 54-72.
- Van Reenen, Pieter T. & Evert Wattel. 1992. De uitspraak van /s/ en /z/ voor klinker in het Nederlands: zes eeuwen variatie. In Bennis Hans & Jan W. de Vries (eds.), *De Binnenbouw van het Nederlands, Een bundel artikelen voor Piet Paardekooper*, 291-309. Dordrecht: ICG Publications.
- Willis, Clodius. 1972. Perception of vowel phonemes in Fort Erie, Ontario, Canada, and Buffalo, New York: An application of synthetic vowel categorization tests to dialectology. *Journal of Speech and Hearing Research* 15. 246-255.