



Anne-France Pinget*

Individual differences in phonetic imitation and their role in sound change

<https://doi.org/10.1515/phon-2022-2026>

Published online November 17, 2022

Abstract: This paper explores the possibility that the spread of sound change within a community correlates with individual differences in imitation capacities. The devoicing of labiodental fricatives in Dutch serves as a case study of an ongoing sound change showing regional and individual variation. The imitation capacities of Dutch speakers born and raised in five regions of the Dutch language area were investigated in a forced imitation task (Study 2) and a spontaneous imitation task (Study 3), and compared to baseline productions (Study 1) of the variable undergoing sound change. Results showed that the leaders of sound change in each region were significantly less accurate in imitating model talkers – when they were instructed to – than conservative speakers, but they were more inclined to spontaneously imitate talkers. These insights are discussed in view of the literature on different types and measures of imitation capacities, on the actors of sound change and the two apparently paradoxical features of the language system: its stability and its potential for sound change.

Keywords: individual differences; laboratory phonology; leaders; phonetic imitation; sound change

1 Introduction

1.1 Phonetic imitation and its role in sound change

Phonetic imitation is the process by which speakers tend to make their speech more similar sounding to that of the speaker they are interacting with. This process has sometimes also been called *phonetic convergence*, *alignment* or *accommodation* depending on the specific subfield of linguistics or theoretical paradigm, with sometimes subtle differences in what is precisely meant (see also Section 1.3.). For the sake of this study, we assume that these terms refer to

*Corresponding author: Anne-France Pinget, Utrecht Institute of Linguistics OTS, Utrecht University, Trans 10, 3512 JK Utrecht, The Netherlands; and Fryske Akademy, Leeuwarden, The Netherlands, E-mail: a.c.h.Pinget@uu.nl

the same process and use the term phonetic imitation. The idea that phonetic imitation plays a central role in the process of sound change can be traced back to the end of the nineteenth century (e.g., Paul 1880; Sievers 1901). It has been suggested that phonetic imitation is the seed for sound change, but also the mechanism (or one of the mechanisms) by which it is spread (e.g., Delvaux and Soquet 2007; Garrett and Johnson 2013; Pardo 2006; Trudgill 2004, 2008).

In the first case, imitation is viewed as the mechanism by which variation can result in sound change. Speakers can and do imitate, but they do so quite imperfectly (e.g., Babel 2012; Pardo 2006), they only approximate other speakers' productions, which results in mutations of the signal and provide the innovations themselves. Imperfect imitation is often investigated in relation to language acquisition, with children as the main actors of change. Children, however, cannot be the only actors of change, since there is ample evidence that adult speakers participate in ongoing changes taking place in their community (e.g., Harrington et al. 2000; Sankoff and Blondeau 2007). In short, if imitation – either by children during acquisition or adults throughout their life – is the seed for linguistic change in general, phonetic imitation is the seed for a specific kind of linguistic change: sound change.

In the second case, phonetic accommodation is considered as one of the main mechanisms by which a change is disseminated through a community (Harrington and Schiel 2017; Trudgill 1986, 2008) as interacting listeners and talkers accommodate to each other's speech patterns. Under the change-by-accommodation model (Auer and Hinskens 2005; Niedzielski and Giles 1996; Trudgill 1986), new variants are spread through the automatic adjustment of phonetic properties in response to an interlocutor. This model consists of three stages: (1) a short-term accommodation in individual interactions, (2) a long-term accommodation resulting in permanent changes in the speech of individual speakers, and (3) the spread of new variants throughout the community. Based on the transition from short-term temporary accommodation to long-term permanent accommodation, this hypothesis manages to relate 'change at the level of the community to variable use in verbal interaction' (Auer and Hinskens 2005: 356). This idea is rooted in the social psychological framework of Communication Accommodation Theory (CAT) (Giles and Smith 1979; Giles et al. 1973), where imitation behavior was considered as social acts that talkers use to modulate social distances in communication.

Delvaux and Soquet (2007) noted that these two views are not contradictory. Indeed, they depart from the observation that phonetic imitation occurring at the inter-individual level has to account for two apparently paradoxical features of the language system: the stability of phonetic realizations within a speech community and their potential for sound change. Delvaux and Soquet (2007)

proposed that – when interacting – speakers and listeners automatically tend to converge and achieve a consensus on the acoustic-phonetic level, resulting in a hybrid realization. At the community level, the accumulation of these consensuses results in a particular set of variants characteristic for a particular language variety. Moreover, these consensuses are constantly renewed and the corresponding mental representations are updated. Sound change can then be defined as a consensus within a community to use a particular phonetic variant that is different from the one achieved previously (Delvaux and Soquet 2007). In this way, phonetic imitation could be both the cause of sound change and the mechanism of its spread.

More generally, it had widely been assumed that sound change indeed relies on two distinct events that occur sequentially: first, the creation of a new variant in speech variation (called *innovation* or *actuation* (Weinreich et al. 1968)), and second its propagation, the spread in the speech community. This conceptualization of sound change in terms of a two-step process of variation and selection draws inspiration from biological evolution (Yu 2013). As pointed out by Labov (1994) however, we do not necessarily need to distinguish between these two processes, as it is peculiar to think of sound change relying on two distinct activities: a sort of creative act of deviating from the existing use (of which it is unclear whether it is even observable) (Croft 2000), followed by the imitation of this innovation by other speakers. We assume that sound change is much more of an iterative process at the individual level in which minimal changes incrementally accumulate in a speaker's system every time he speaks to a listener (as an intrinsic result of phonetic imitation). Under such a theoretical assumption, sound change might be defined as a purely synchronic process, residing in the variability inherent to production and perception. The propagation of a change is mostly observable when comparing groups of speakers (stratified by age for instance or any other broad social category). A sound change is 'actuated' every time again in new speakers and new communities. Diachronic change is then merely a series of incremental acts of actuation in a defined direction.

1.2 The individual versus the group level

Based on these insights about the role of imitation in change, it is interesting to examine how and why sound change spreads through space. If we consider a language area to consist of smaller regional communities, the consensuses achieved in a region might logically slightly differ from the consensuses in other regions, since there is more inter-personal contact within than between regions. As these consensuses are constantly renewed and updated, sound changes can spread through the language area.

Dialectologists have mostly been concerned with this geographical dimension of sound change, and thus with the group level. Sociolinguists who study change over time have also been interested in the question of who leads language change (Labov 2001; Tamminga 2021). Their research has extensively shown that the advancement of a sound change is not uniform in one single community (e.g., Harrington and Schiel 2017; Milroy 2002; Milroy and Milroy 1985). Even in a small local community sound change is not equally advanced in each member, leading to a puzzling question: where do these individual differences come from? Various studies have aimed to identify speakers who are at the forefront of ongoing sound change, those often called *leaders* of change, *innovators* or *early adopters* (with some subtle differences between these terms). To date, many characteristics of the leaders of change have been revealed, most of them related to social factors in Western societies, such as network density and size (e.g., Lev-Ari 2018; Milroy 1987), gender (e.g., Eckert and McConnell-Ginet 2003; Labov 1990, 2001) or social class (e.g., Ash 2002; Chambers 2002; Labov 2001). More recently, the possibility that this lack of uniformity in the spread of a change within a single region or community is related to individual differences in linguistic, social and cognitive aspects has been proposed. Baker et al. (2011) looked at /s/-retraction in American English and Beddor (2012) at coarticulation patterns. They proposed that individual speakers differ in the extent to which they respectively retract versus coarticulate and therefore do not all participate to the same extent in the sound change. Garrett and Johnson (2013) proposed that some speakers are more likely to index linguistic differences with social meaning than others and that this difference in indexicality is a driving force in sound change. Yu (2013) explored individual variability in cognitive processing in relationship to the spread of sound change throughout a community. He found that the personality and social profiles of some individuals (more extroverted and agreeable) can predict the spread of sound change in their networks. Tamminga (2021) undertook a preliminary investigation of leadership and personality traits in order to predict the spread of a set of sound changes in Philadelphia English. Based on her results, she showed some pessimism towards the idea that such traits can predict individual differences in change advancement. The current study primarily aims to bring these attempts further by exploring to extent to which individual differences in change advancement can be linked to phonetic imitation capacities.

1.3 The types of phonetic imitation

In sociophonetic research different paradigms have been developed to examine phonetic imitation (Pardo 2013). The shadowing paradigm was first introduced

into the field by Goldinger (1998) and subsequently used in numerous imitation studies. In a shadowing study, participants typically listen to and repeat isolated words or sentences. Results within this paradigm have repeatedly showed that subjects shift their speech production in the direction of speech they are asked to repeat. These shifts are constrained by linguistic factors (e.g., Nielsen 2011) and modulated by language-external factors, which result in significant variability between speakers in the extent and the directionality of shifts (convergence vs. divergence patterns). Factors such as attitude toward the interlocutor (Abrego-Collier et al. 2011), gender both of the speaker and listener (Babel 2010; Namy et al. 2002), personality traits (Yu et al. 2013) and cognitive load of the task (Abel and Babel 2017) and have been shown to constrain these phonetic imitation patterns.

Additionally, efforts were made to examine phonetic imitation involved conversational interaction, a setting where imitation is traditionally thought to be more relevant than within the non-interactive shadowing task (see Communication Accommodation Theory in Giles and Smith 1979 and Giles et al. 1973). Delvaux and Soquet (2007) designed a task in the lab with a higher degree of social interaction than in previous studies. They observed what they called *deliberate imitation*: when simply exposed to the other regiolect, speakers of one regiolect produced vowels that were significantly different from their typical realizations, and significantly closer to the other regiolect. Pardo et al. (2018) compared phonetic convergence in conversational interaction and in a non-interactive speech shadowing task and found that patterns of phonetic convergence highly differed across these settings. Sonderegger et al. (2017) managed to investigate imitation outside the lab by looking at the recordings of the reality television show *Big Brother*. Speakers participating in the show were locked up together in an isolated house for three months. Their methodology also allowed to focus not only imitation in the short term, but to explicitly look at the medium term (i.e., a few month). They looked at five phonetic variables and found that small daily accent fluctuations were ubiquitous, while more persistent accent changes occurred only in a minority of speakers.

These attempts to examine phonetic imitation in more spontaneous, non-laboratory and interactional settings focused on language variation. Yet, it remains crucial to compare synchronic variation patterns to the specific situation of sound change. Especially in the context of sound change, it is interesting to investigate not only whether people are willing to imitate, but also whether they are capable of imitating. By definition, sound change means that the present and future realizations of a variable are different than the original, past realizations. Some speakers might at some point in the change lose the ability to produce a sound in a certain way. Therefore, it appears important in the case of sound

change to consider both the (sociolinguistic) willingness or readiness and the (phonetic) articulatory or auditory capacity to imitate.

To the best of our knowledge, only Dufour and Nguyen (2013) have investigated what we can called *forced* and *spontaneous* imitation separately. Using a stable vowel variable in French, they compared the phonetic convergence effect observed in – on the one hand – an imitation task in which participants were explicitly instructed to imitate the productions they were exposed to (forced imitation), and on the other hand – a shadowing task which was meant to trigger unintentional imitation (i.e., spontaneous imitation). They found that the phonetic convergence effect was greater when participants intentionally imitated the speaker's productions than in the shadowing task, and that both types of instruction led to the same degree of convergence in a post-exposure task.

The goal of this study is to answer the question to what extent these two different types of imitation play a role in the case of a sound change, and what their respective contribution is to the spread of a change. We hypothesized that the speakers who tend to easily imitate spontaneously might introduce the change into their own community, while speakers who fail to imitate when instructed to, take a conservative position in the change: they keep on reproducing the existing speech patterns in a precise way, and in doing so, reinforce the stability of phonetic realizations within the community. In conclusion, we expect that there is a correlation between speakers' innovativeness in sound change and their spontaneous imitation capacities on the one hand, and between speakers' conservatism in the change and their forced imitation capacities in the other hand.

1.4 This study

In the current paper, we explore individual differences in imitation capacities for a Dutch variable that is involved in the process of sound change. The devoicing of labiodental fricative /v/ is a well-attested ongoing sound change in the Dutch language area, showing regional and individual variation (see Section 1.5.). Those patterns are investigated in a baseline production task (Experiment 1). Subsequently, we explore imitation capacities in a forced imitation task (Experiment 2) and a spontaneous imitation task (Experiment 3) and the extent to which the imitation results relate to individual differences in change advancement. Unlike the order in which the data are presented in this paper, participants conducted the tasks in the following order: the baseline production experiment, the spontaneous imitation experiment and finally the forced imitation experiment. The order permitted to avoid that the forced imitation task, being inherently meta-linguistic, influenced the spontaneous imitation task.

1.5 A sound change in progress: the devoicing of labiodental fricatives in Dutch

Standard Dutch is traditionally described as having a phonological distinction between voiced and voiceless fricatives. The major cue for the voiced/voiceless distinction is the presence or absence of vocal fold vibration in the fricative (Slis and Cohen 1969). During the last decades, it has been frequently observed that word-initial voiced fricatives in standard Dutch are increasingly produced as voiceless (Cassier and Van de Craen 1986; Cohen et al. 1961; Gussenhoven 1999; Hamann and Sennema 2005; Kissine et al. 2003; 2005; Mees and Collins 1982; Van de Velde 1996; Van de Velde et al. 1996; van der Wal et al. 1992).

Regional differences are observed 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 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 /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 conclusion, this sound change shows significant regional variation and can be considered as advanced, but not completed at the moment we are conducting this study. Pinget et al. (2020) showed that the regional differences in the amount of devoicing in fricatives matched the patterns described in previous studies and appeared to be even further advanced than reported two decades ago by Kissine et al. (2003, 2005). They also showed that there was a clear link between the production and perception systems undergoing sound change.

2 Experiment 1: baseline production

In this section, we report the results of the baseline experiment (also reported in Pinget et al. 2020) which was designed to investigate the individual and region production patterns of labiodental fricatives /v/ and /f/. First, the method is described in Section 2.1. Production data are presented in Section 2.2, showing regional stratification and individual differences in this sound change.

2.1 Method

2.1.1 Regions and participants

Based on the studies described in Section 1.5, five regions within the Dutch language area were chosen to reflect different stages of fricative devoicing: West-Flanders (WF), Flemish-Brabant (FB), Netherlands Limburg (LI), South-Holland (SH) and Groningen (GR). These regions were selected to represent smaller communities within the larger Dutch language area. They are represented on the map in Figure 1.

West-Flanders (WF) is a peripheral region to the West of Flanders along the North Sea. The chosen area is situated around the towns of Kortrijk and Roeselare. This region is known to be the most conservative in terms of fricative devoicing (Kissine et al. 2003, 2005). 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. Weak patterns of fricative devoicing have been found in this region (Kissine et al. 2003, 2005). Netherlands Limburg (LI) is a geographically peripheral region situated in the South of the Netherlands, stretching from Venlo to Maastricht. The fricative devoicing process is weak to moderate in this region (Kissine et al. 2003, 2005). South-Holland (SH) is part of the

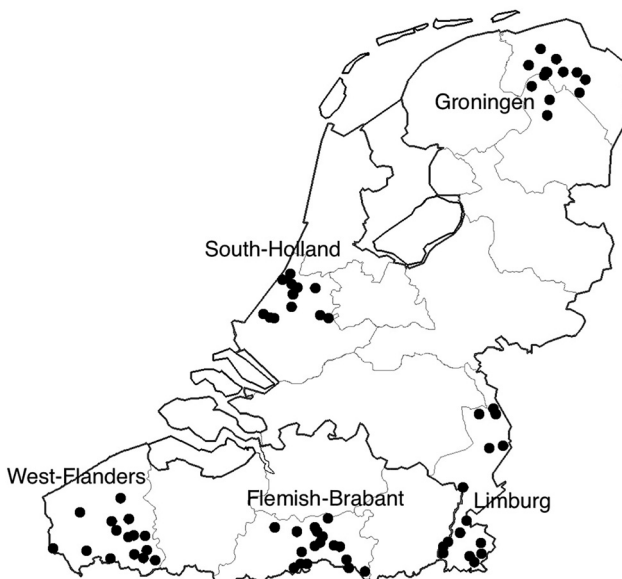


Figure 1: Map of the Dutch language area (The Netherlands and Flanders only) and of the five selected regions. Each dot represents the origin of one or more participants ($n = 20$ per region).

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. Production studies have shown strong devoicing of fricatives in this region (Kissine et al. 2003, 2005) and it is often considered as the area from which this phenomenon is spreading (Van de Velde et al. 1996). The Groningen region (GR) is situated in the North of the Netherlands and is centered around the cities of Groningen and Assen. It is known to be a region where the voiced/voiceless fricative contrast has almost completely faded, resulting in a merger (Kissine et al. 2003, 2005).

The participants were one hundred native speakers of Dutch born and raised in these five regions. Of each region, 10 males and 10 females took part in the production experiments. Participants were all highly educated young adults aged between 18 and 28 years (mean = 22.03 years). All participants were attending or recently graduated from a university or college, and were fluent speakers of colloquial standard Dutch. No participant reported having any hearing or speaking problems.

2.1.2 Data collection

Participants took part in five different production tasks in standard Dutch: a word reading, a carrier sentence reading, a sentence reading, a semi-spontaneous speech, and a spontaneous speech task. These tasks differed in the amount of attention paid to speech and were intended to elicit the range of phonetic realizations for each individual speaker within the standard variety.

In the word reading task, each participant read a list of words presented in isolation on the screen in a randomized order, including words beginning with labiodental voiced ($n = 20$) and voiceless fricatives ($n = 19$), and fillers ($n = 82$). In the carrier sentence reading, Dutch non-words beginning with voiced fricatives were produced in the frame “Ik neem de ___” [I take the ___] ($n = 9$ for voiced fricatives and $n = 9$ for voiceless fricatives). Next, each participant read a set of declarative sentences in which Dutch words starting with voiced ($n = 14$) and voiceless fricatives ($n = 14$) were elicited. Semi-spontaneous productions of fricatives 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. 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 min. Twenty fricative tokens ($n = 10$ for voiced fricatives and $n = 10$ for voiceless fricatives) from the semi-spontaneous production task and the first ten tokens from the spontaneous production task were analyzed for each speaker, all of them in onset position

and preceded by a vowel. In all five tasks, a maximum of 58 (v) tokens and 57 (f) tokens per participant was analyzed.

All five production tasks 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.

2.1.3 Phonetic measures

All recordings were sampled at 48 kHz, 24 bits. Labiodental fricatives realizations were segmented based on their centre of gravity (Gordon et al. 2002; Jassem 1979; van Son and Pols 1996), following a segmentation protocol for Dutch (van Son 2000). The center of gravity (CoG) was calculated in the domain of 0–16,000 Hz without pre-emphasis of the signal prior to weighting. The onset of fricatives was manually determined based on the start of noise (rising CoG values) and the offset of fricatives by the end of the noise (falling CoG values).

Following Kissine et al. (2003), VOICING was calculated by measuring the fundamental frequency (f0) (in Hertz) with intervals of 10 ms in the fricative segment. The presence of voicing was assessed between 50 and 400 Hz. To compute a voicing score, the number of measurements with presence of f0 (categorically coded) was divided by the total number of measurements and multiplied by 100. The resulting voicing percentage indicates the proportion of voicing in each fricative and ranges from 0% (no voicing throughout the fricative) to 100% (voicing throughout the entire fricative). This measure of fricative voicing will be reported throughout the paper as the main phonetic correlate in the voiced-voiceless contrast. In previous studies on the same production samples (Pinget 2015; Pinget et al. 2020; Pinget and Quené 2021), duration and F0 at adjacent vowel onset were investigated as secondary cues. For both secondary cues, it turned out that they are used in production to realize the contrast (/f/'s are significantly longer than /v/'s, and F0 at the onset of a vowel following /f/ is significantly higher than at the onset of vowel following /v/), but both duration and f0 differences are disappearing as the sound change is proceeding. At the word, individual and regional levels, it was shown that the less voicing in the fricatives, the longer the duration, and the higher the F0 at onset. In the current study, we focus mainly on voicing as the main cue, but will return to the discussion about how the three cues interact in Experiment 2.

All observations greater than four standard deviations from the mean voicing were considered as outliers and removed from the data. In this way, the procedure

managed to remove extremely deviant observations, errors in measurements and speech errors. A total of 10261 fricatives was analyzed (4794 /f/ and 5467 /v/).

2.2 Results

Voicing measures for /v/ and /f/ (in %) split up by region are presented in Figure 2. First, we observed that voicing measurements for /f/ are very stable across regions and individuals, while regional and individual differences in the realization of /v/ are very large. As expected, West-Flemish participants produced /v/ with the highest degree of voicing (mean = 56.64% of voicing). Flemish-Brabant (mean = 49.84%), Limburgian (mean = 43.02%) and South-Hollandish voiced fricatives (mean = 29.64%) show gradually more devoicing in this order. Groningen participants produced the most devoiced /v/'s (mean = 19.85%) with values not different from /f/'s: they have a merged /v/ and /f/ production. There are however two speakers in Groningen who produced /v/ with much more voicing than the other speakers of the same region.

Altogether, these results confirmed previous work on the devoicing of labiodental fricatives: it is an advanced sound change showing regional stratification. The devoicing is resulting in a merger: /v/ is merging into /f/, and not the other around. This merger is the most advanced in the regions of South-Holland and Groningen.

Furthermore, there are large individual differences in the production of /v/ within each region. In other words, the spread of the change is not uniform within

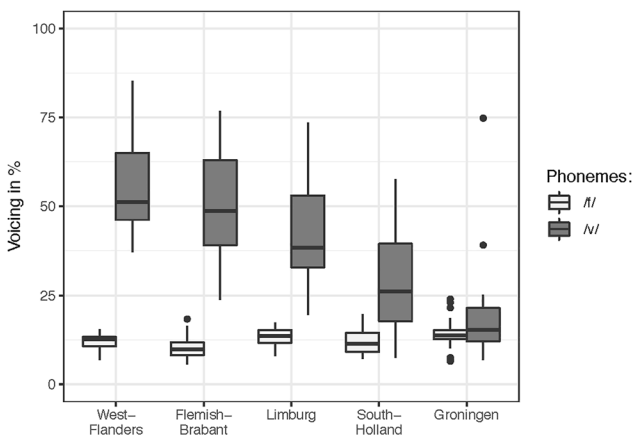


Figure 2: Boxplot of voicing measures (in %) for /v/ (in grey) and /f/ (in white) in the aggregated data (one value represents one speaker), split up by region.

the different regions. The amount of voicing in /v/ realizations were fit with a linear mixed effects model using the *lme4* package in R (Bates et al. 2015). The model intends to control the effects of a range of linguistic and social predictors that are not the target of this individual differences investigation, and which were included as fixed-effects predictors: (1) speech style (i.e., word reading, carrier sentence reading, sentence reading, semi-spontaneous speech, spontaneous speech), (2) region (West-Flanders, Flemish-Brabant, Limburg, South-Holland, Groningen), (3) speaker gender (male, female) and (4) speaker age (at time of interview). The model also included by-word and by-speaker random intercepts.¹ Following a procedure previously used by e.g., Drager and Hay (2012), Voeten (2021) and Tamminga (2021), we take the by-speaker random intercepts as a measure of how innovative or conservative a speaker is in comparison to the other speakers in the dataset within their own region, controlling for the other social and linguistic factors. This procedure provides a determination of every speaker's place on the innovativeness continuum. When random intercepts are positive, speakers are more conservative in the change than their peers. Producing more voicing is thus conservative in the case of a devoicing process. A zero intercept means that the speaker's average voicing production is equal to the regional mean. A negative intercept indicates that the speaker is more innovative than the peers. This measure of speakers' innovativeness will in the subsequent part of the paper be used to compare the advancement in the sound change with imitation capacities (forced imitation in Experiment 2 and spontaneous imitation in Experiment 3).

3 Experiment 2: forced imitation

In Experiment 1, a forced imitation experiment was conducted. This experiment aimed to explore participants' capacities to produce the whole range of voicing in labiodental fricatives and a phonetically accurate imitation of these consonants when they instructed to.

3.1 Method

3.1.1 Participants

The same participants as described in Section 2.1.1 took part in this experiment.

¹ The syntax of the fitted model was: `model1 <- lmer(voicing_v ~ SpeechStyle + SpeakerRegion + SpeakerGender + SpeakerAge + (1 | Word)+(1 | Participant), ProductionData)`.

3.1.2 Model speaker and stimuli

A male speaker of South-Holland region (25 years old, trained phonetician) served as the model talker. He produced /vi/ and /fi/ syllables in carrier sentences. His productions were digitally recorded with a sample frequency of 44.1 kHz in a sound-attenuated cabin and subsequently used to create a /vi/-/fi/ speech continuum, generated by manipulating voicing.

The fricatives of the source recordings were extracted from their original context and used as the extremes of the continuum along the voicing dimension (with respectively 0 and 100% voicing). Five steps were generated by spectral linear interpolation, using the PSOLA (Pitch-Synchronous-Linear-Overlap-and-Add) algorithm of Praat (Boersma and Weenink 2022) (based on the script of Mitterer 2009). The interpolation provided in-between realizations characterized by approximately 25, 50, and 75% of voicing. The five steps of the continuum were originally also manipulated for duration with five steps (60 ms, 94 ms, 128 ms, 162 ms, 196 ms), resulting in a bi-dimensional continuum with 25 stimuli in total (see also Pinget et al. 2020 where a similar, but larger bi-dimensional continuum was used in the perceptual identification task). The following vowel had a constant duration of 110 ms and its f_0 contour was flattened around 135 Hz through a PSOLA pitch manipulation. In this manner, f_0 contour in the vowel was in fact removed as a potential additional cue. All these manipulations resulted in a 25-steps /vi/-/fi/ continuum.

3.1.3 Procedure

Participants were seated in a sound-attenuated booth and listened to the 25 /vi/ and /fi/ syllables via Beyerdynamic DT 250 headphones. They were instructed to imitate the speaker as accurately as possible after each stimulus, and to embed their imitation in a carrier sentence. The carrier sentence was 'ik neem de ___' [I take the ___]. The sentences were recorded through an AKG C420 head-mounted microphone.

Each participant completed a practice block of six trials and the experimental session with five repetitions of the five continuum steps presented in a randomized order ($n = 25$ per participant). The task was self-paced, required a high attention level and lasted for approximately 4 min.

3.1.4 Phonetic measures

All recordings were sampled at 48 kHz (24 bits), segmented and labeled as described in Section 2.1.3. Imitation was measured acoustically by examining

voicing in the fricatives measured as described in Section 2.1.3. In addition, we measured two additional potential cues: (1) the fricative total duration in milliseconds, and (2) F0 at the onset of the following vowel. F0 measurements were taken using Praat (Boersma and Weenink 2022) at 10 equidistant time points within the vowels. F0 was assessed between 100 and 500 Hz for females, and between 75 and 300 Hz for males. F0 measures were examined visually and checked by hand. For each individual speaker, all f0 values subsequently were averaged, yielding a speaker-specific f0 centroid in Hz. All f0 measures of a speaker were subsequently expressed in semitones relative to this speaker-specific f0 centroid (see Shultz et al. 2012).

3.2 Results

3.2.1 Inspection of all three phonetic cues at the regional level

First, all three phonetic measures (voicing, duration and f0 patterns in following vowel) were investigated in the imitated fricatives. These results are presented in Appendix 1, split up by (Figure A) regions. A detailed discussion of all effects goes beyond the scope of this paper, but we describe here the results which are necessary for the discussion of the results at the individual level in the next Section (3.2.2.). First, voicing, duration and f0 patterns are compared between the target (manipulated in the bi-dimensional continuum) and imitated fricatives (produced by the participants): the more voicing in the target fricatives, the more voicing in the imitated fricatives. Similarly, the longer the target fricatives, the longer the imitated fricatives, but overall imitated fricatives were much longer than the presented targets. This could mean that the role of duration is enhanced in the task due to its nature. Based on the analysis of these two cues, it appeared that the forced imitation task successfully triggered imitated fricatives which varies along the voicing and duration continuum in a way that approximates the target stimuli. Additionally, we saw that the more voicing in the target fricatives, the higher F0 in the following vowel tended to start. This effect is surprising since F0 was flattened in the target stimuli. This seems to point at the facts that: (1) F0 in the following vowel is not necessary to perceive the fricative contrast in the presence of other cues (voicing and duration), and yet (2) F0 is indeed used as a cue highly correlating with voicing in the produced imitations, showing that the presence of this cue is probably the result of an automatic and articulatory-motivated process (see e.g., Halle and Stevens 1971; Hanson 2009). Finally, all these cue effects were visible in all five regions, meaning that they were stable even in the context of strong devoicing.

3.2.2 The use of voicing and its relationship to the advancement of the sound change within individual speakers

In this section, we focus on the use of voicing as a cue in the forced imitation task and how the use of this cue relates to the advancement of the sound change at the individual level. First, the range in voicing that individual speaker produced in the imitation task was investigated. For each participant, the voicing range between the most voiced and most voiceless realizations in the imitation task was computed, and was expressed between 0% (no range) and 100% (full voicing range). Most participants were able to produce a voicing range of (nearly) 100%. However, a few participants did not achieve high voicing ranges. Six participants showed a voicing range around 50% and four participants showed ranges lower than 25%. These participants mostly failed to produce enough voicing when imitating the voiced stimuli.

Secondly, the accuracy of the forced imitation was examined. For each trial in the forced imitation task, the amount of voicing in the target fricatives produced by the model talker was subtracted from the amount of voicing produced by the speaker. This measure is called voicing accuracy. Whenever voicing accuracy is positive, there is overshoot: the speakers produced too much voicing compared to the model talker they were instructed to imitate. A negative score indicates voicing undershoot: the speaker failed to produce enough voicing compared to the token produced by model talker. Zero in voicing accuracy represents a perfect imitation.

We fitted a linear mixed effects model using the *lme4* package in R (Bates et al. 2015) with the speakers' degree of innovativeness (see Section 2.2) and the total voicing range (i.e., across the whole forced imitation task) as predictors for the voicing accuracy. The model also included by-speaker and by-word random intercepts.² There were significant main effects of the degree of innovativeness ($\beta = 0.379$, $t = 3.679$) and of the range innovativeness ($\beta = 0.416$, $t = 5.127$), reflecting the facts that more innovative in the change, the more speakers undershoot during forced imitation and the smaller the voicing range they can produce.³

These effects are visualized in Figure 3, with on the x -axis the speaker's degree of innovativeness within their own region (based on the baseline production task)

² The syntax of the fitted model was: `model2 <- lmer(VoicingAccuracy ~ SpeakerInnovat + TotalRange_v + (1 | Trial)+(1| Participant), ForcedImitationData)`.

³ The same model was run on a subset of the participants (only the participants who took part in the spontaneous imitation task, i.e. participants from the regions of West-Flanders, Flemish-Brabant and Limburg ($N = 60$)). The main effect of the degree of innovativeness as a predictor for voicing accuracy computed on the subset turned out very similar to the same main effect computed on the whole sample ($\beta = 0.530$, $t = 4.263$).

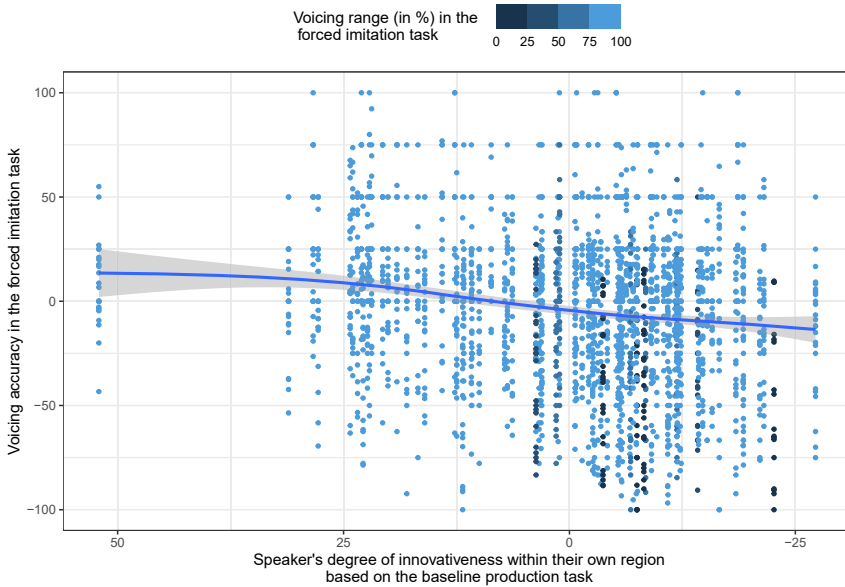


Figure 3: Scatterplot of voicing accuracy in the forced imitation task as a function of the speaker's degree of innovativeness within their own region (based on the baseline production task). Each symbol represents a trial in the forced imitation task ($n = 2468$), the gradient color indicates the total voicing range.

(see Section 2.2) and the y -axis the voicing accuracy in the forced imitation task, split up by voicing range in gradient color.

3.3 Intermediate discussion

The first imitation experiment was a forced imitation task in which participants were explicitly instructed to imitate a model talker. This experiment was aimed at measuring participants' capacities to produce the full range in voicing in labiodental fricatives, and at producing a phonetically accurate imitation of these consonants.

First, participants' capacity to produce the whole range of voicing in labiodental fricatives was investigated. In the context of (strong) devoicing, it was examined whether individual speakers were still able to produce fully voiced variants when they were instructed to, even if they had lost the contrast in their daily productions. It was shown that the large majority of participants – even in regions where devoicing is advanced (such as South-Holland and Groningen) –

was able to produce the entire voicing range. Most speakers appeared to have maintained to some extent the articulatory ability to produce /v/ when explicitly instructed to. Yet, around 10% of the participants failed to produce full ranges. This might be due either to a diminished sensitivity in perception, or to an articulatory difficulty to produce the contrast, or both at the same time (see discussion in Section 6).

Second, it was examined whether the accuracy of the forced imitation (i.e., the extent to which imitated productions resembled the targets) can predict the advancement of speakers in their own region. It turned out that the more speakers are advanced in their own region, the more undershoot of phonetic imitation of the model speaker. Leaders of change within a community seemed to be worse at accurately imitating fricatives, while more conservative speakers showed high phonetic imitation accuracy (around 0 in voicing accuracy score or a slight voicing overshoot). In conclusion, the forced imitation task demonstrated how far participants can still push their production of labiodental fricatives when they are instructed to, and gave insight into how motoric imitation capacities are related to the spread of change. Whether participants are able to use these motoric capacities in social interaction is the question raised in the second experiment: a spontaneous imitation task.

4 Experiment 3: spontaneous imitation

The third experiment was meant to elicit spontaneous imitation. This is thus imitation patterns that participants are likely not aware of, and that they produce spontaneously without being instructed to do so. The methodology was largely based on Delvaux and Soquet (2007): the aim was to test whether – when exposed to model talkers with devoiced /v/ – speakers would spontaneously produce fricatives that are more devoiced than their baseline realizations, thus go along with (the direction of) the sound change.

4.1 Method

4.1.1 Participants

A subset of the original participant sample (described in Section 2.1.1.) was selected for the spontaneous imitation task. Only the participants from the regions of West-Flanders, Flemish-Brabant and Limburg ($N = 60$) took part, as their

average realization of /v/ was more voiced than the realizations of the model talkers. In other words, the average /v/ voicing in those regions significantly differed from the devoiced productions by the model talkers in the spontaneous imitation task, so that participants actually have something to converge to.

4.1.2 Model speakers and stimuli

Two female speakers served as model talkers. Model speaker 1 came from Flanders (26 years old, from Antwerp) and model speaker 2 came from the Netherlands (32 years old, from North Brabant). Both model speakers read 36 target sentences digitally recorded with a sample frequency of 44.1 kHz in a sound-attenuated cabin of the lab of Utrecht University.

All target sentences were of the type: *de (object) gaat in de (container)* ‘the (object) goes into the (container)’. The objects were 36 easily recognizable common objects (i.e., a flower, a fork, a banana, a paperclip etc.). There was no particular linguistic or external constraint on these words, except that they had to refer to objects of a size small enough to be put into the containers. The container words were *vuilnisbak* ‘trash bin’ containing the initial voiced labiodental fricative as target variable and *boekentas* ‘schoolbag’, as distracter. Both were low frequency words, since Goldinger (1998) showed that low-frequency words tend to show more imitation than high frequency words. Both containers were three-syllabic words preceded by a schwa in the carrier sentence.

One highly devoiced realization of the voiced fricative in *vuilnisbak* (‘trash bin’) for each model talker (model speaker 1: 36% of voicing and model speaker 2: 23%) was selected and subsequently concatenated with the 36 recorded object complements (f.i., *de aarbei gaat in de* ‘the strawberry goes to the’). These realizations of the voiced fricative by the model talkers were significantly more devoiced than the regional mean in West-Flanders, Flemish-Brabant and Limburg (see Section 2.2.). No audible disturbance was present in the concatenated stimuli. In this way, 36 stimulus sentences were obtained for both model speakers, each of them containing one of the naturally produced 36 different objects with the same realization of the devoiced fricative.

4.1.3 Procedure

The experiment took the form of a card game played by three players: the participant and the two model talkers, called Anna (A) and Lisa (L) in the experiment. Screen shots of the experiment are provided in Figure B in Appendix 2. In each trial,

a card representing one of the 36 objects appeared on the screen, together with an arrow pointing towards either the schoolbag or the trash bin. When their turn came, players had to orally formulate the association they saw between the object and the container by pronouncing a full sentence (f.i. *de aardbei gaat in de vuilnisbak* ‘the strawberry goes into the trash bin’). A and L’s voices were played through Beyerdynamic DT 250 headphones. Participant’s (P) sentences were recorded through an AKG C420 head-mounted microphone. The instruction given to participants was to perform the task only when their turn came (every three trials). The name of the player in turn was colored on the screen. The order of the turns was A L P L A P A L P... so that the participant’s turn followed a different model speaker in each trial (see Appendix 2). The expectation based on previous interactional imitation studies (e.g., Babel 2012; Delvaux and Soquet 2007; Dufour and Nguyen 2013) is that participants will (gradually) converge in their productions to the variant (i.e., in this case the devoiced fricative) produced by the model talkers. Alternatively, it might be the case that some participants (gradually) diverge from the model talkers.

A cover story was used to conceal the real purpose of the experiment, so that imitation (if present) remained fully spontaneous. Participants were told that the purpose of the experiment was to assess their memory and attention abilities. They were told that the two other players were other participants in the study. They were instructed to learn the associations between the objects and their container, and were asked in a post-test to recall as many of these associations as possible. The cover story was very successful and participants were at no point aware of the real purpose of the experiment. On the contrary, they were challenged by the task: most of them were eager to know their ‘memory score’ at the end of the experiment. They were told in the debriefing about the real purpose of the experiment.

Each participant completed a practice block of 18 trials. The test phase consisted of 180 trials presented in a randomized order (five repetitions of 36 different objects). Each player produced one third of these trials, i.e., 60 sentences (30 with the target variant in *vuilnisbak* and 30 with the distractor *boekentas*). In this way, 30 realizations of /v/ were obtained per participant.

4.1.4 Phonetic measures

All recordings were sampled, segmented and labeled as described in Section 2.1.3. Voicing in /v/ realizations was measured acoustically as described in Section 2.1.3. Outliers were removed in the same way as in the production data ($n = 21$), so that 1779 fricative realizations remained available for analysis.

4.2 Results

Firstly, the extent to which individuals converge to or diverge from the model talkers was investigated. For each speaker, a linear regression was fitted to the 30 /v/ trials with the order of trials as predictor. In each model, a slope significantly different from 0 indicates a gradual imitation pattern: either convergence when the slope is negative or divergence when the slope is positive slope. Figure C in Appendix 2 shows an example of individual patterns of convergence. Six participants from a total sample of 60 showed a significant slope in the regression models: four slopes were significantly negative with the speakers spontaneously converging towards the model talkers, while two slopes were significantly positive with speakers diverging from the models. For 54 participants, the order of trials as predictor had no significant effect on the amount of produced voicing.

Secondly, we examined to what extent speakers' /v/ realizations were influenced by the model talkers. Importantly, the distances that each speaker has to cover in order to converge towards the model talkers necessarily depends on their baseline production. We therefore subtracted the individual baseline production of /v/ as obtained in Experiment 1 from the /v/ realizations in the spontaneous imitation task at the by-talker and by-trial level. The obtained scores represent the modification in voicing of /v/ induced by the exposure to the model talkers. A positive modification score means that /v/ productions in the spontaneous imitation task are more voiced than the speaker's baseline production. A negative modification score means that /v/ productions in the spontaneous imitation task are less voiced compared to the baseline, that is what we expect as the model talkers have a highly devoiced /v/. Zero means that the speaker's production was not influenced by the model talker. We fitted a linear mixed effects model using the *lme4* package in R (Bates et al. 2015) with the speakers' degree of innovativeness (see Section 2.2.) as predictor of the amount of voicing in the imitated /v/ realizations. The model also included by-speaker and by-word random intercept.⁴ There was a significant main effect of the degree of innovativeness ($\beta = 0.406$, $t = 2.512$), showing that more innovative in the change, the more speakers show a voicing modification: their /v/'s became less voiced compared to the baseline as a result of exposure to model talkers.

This effect is visualized in Figure 4, with on the *x*-axis the speaker's degree of innovativeness within their own region (based on the baseline production task) (see Section 2.2) and on the *y*-axis the voicing modification induced by the model

⁴ The syntax of the fitted model was: `model3 <- lmer(Voicing_v ~ SpeakerInnovat + (1 | Trial)+(1 | Participant), SpontaneousImitationData)`.

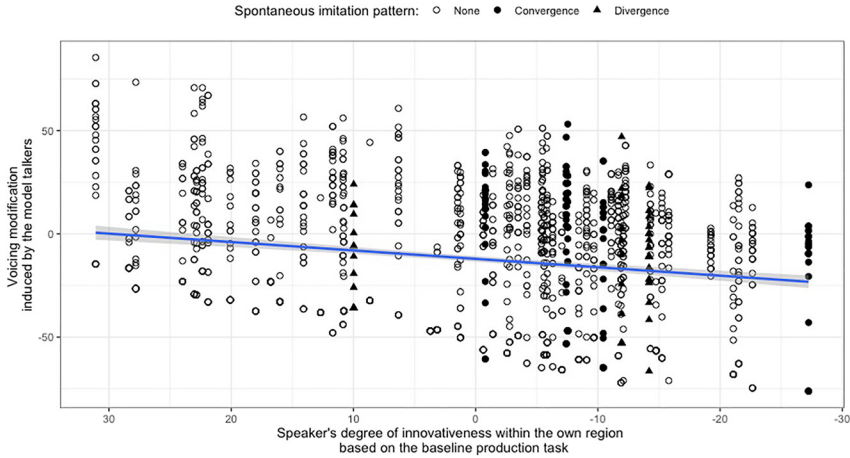


Figure 4: Scatterplot of voicing modification induced by the model talkers in the spontaneous imitation task as a function of the speaker's degree of innovativeness within their own region (based on the baseline production task). Each symbol represents a trial in the spontaneous imitation task with an indication of the general gradual imitation pattern, either no significant pattern (open circle), a pattern of convergence towards the model talkers (filled circle) or a pattern of divergence from the model talkers (filled triangle) ($n = 1779$).

talkers, with different shapes depending on the general gradual imitation pattern over the time course of the experiment.

The fitted line visualizes the main effect found in the model, whereby the most conservative speakers in each region (with a positive degree of innovativeness) showed no voicing modification as compared to their baseline (around 0), while the most innovative speakers (with a negative degree of innovativeness) showed a lower voicing in /v/'s produced in the imitation task compared to their baseline, as a result of exposure to the model talkers with high degree of devoicing. Moreover, we observed that the four speakers who showed a significant tendency to gradual convergence towards the models are all quite innovative in their own region (to the right side of the x-axis), which confirms the relationship between the readiness to imitation and the advancement in the change.

4.3 Intermediate discussion

This spontaneous imitation experiment aimed to investigate participants' readiness to use the devoiced variant of labiodental fricatives in a social interaction.

First, it turned out that a small proportion of participants (10%) showed significant gradual imitation pattern of convergence versus divergence in the course of the experiment. Speakers who gradually converged were all relatively innovative in the change. This small proportion of spontaneous imitators is similar to what was found in similar previous studies (Sonderegger et al. 2017). The assumption was here that the task gradually triggers more imitation with more exposure to the model talkers. However, it only contained 30 trials of the target variables (apart from the practice trials and the distractors). It cannot be excluded that some speakers who did not imitate in the task would have needed more time and exposure (i.e. more trials) to converge to the model talkers.

Furthermore, it was examined whether the modification in voicing of /v/ in the imitation task compared the baseline could predict the position of speakers in their own region. It turned out that the more advanced speakers in their own region, the more they were inclined to imitate the model talkers. Leaders of change within a community seemed to be more ‘willing’ to deviate (even more than what they normally do), while more conservative speakers are less inclined to spontaneously imitate model speakers producing the new variant. In conclusion, the task succeeded in triggering – for most speakers – phonetic convergence (at least at the token level) through simple exposure to two model speakers, even if the social situation was minimally interactive (as compared to usual social situations in which one plays card games).

5 Discussion and conclusion

In this paper, we reported on three types of experimental data collected from the same pool of participants: production data, forced imitation and spontaneous imitation data. In the forced imitation task, participants were explicitly instructed to imitate target stimuli, whereas in the spontaneous imitation task, participants were at no point aware that they were taking part in an imitation task. These two experiments triggered different types of phonetic imitation which were compared to the baseline production data.

The results from the production tasks (Experiment 1) have confirmed that the devoicing of labiodental fricatives is an advanced sound change in Dutch (even more advanced than reported a decade ago by e.g. Kissine et al. (2003, 2005), and that this change shows both regional and individual differences. Apart from spreading across regions, this sound change is also gradually spreading within regions. Based on the random part of a mixed-effects regression model of the

baseline /v/ productions, we computed every speaker's position on the innovativeness continuum within their own region. It allows to distinguish – within each region – between the leaders of the change and more conservative speakers. The goal of the two subsequent imitation tasks was to explore the possibility that the spread of the change within communities could be explained by individual differences in imitation capacities.

The forced imitation task (Experiment 2) provided insight into speaker's phonetic control when producing fricatives despite the ongoing sound change. By explicitly instructing participants to imitate fricatives produced by a model talker, their maximal range of production was investigated. Most participants managed to produce the whole range of fricative voicing. The accuracy of their forced imitations, measured as the distance between the target and the imitation productions, however showed large individual differences. Leaders of the change in each region tended to be less accurate in phonetic imitation than the more conservative speakers: they undershoot the targets more often (i.e., produced in general with less voicing than the target). The best forced imitators were quite average in their own region or at the conservative side in the change. While this correlation between change advancement and imitation capacities clearly emerged in the data, we are facing a chicken-and-egg situation. The question is whether the diminished capacity to imitate causes speakers to be further in the change, or whether their position in the change causes them to poorly imitate when instructed to. Both potential causal relationships will now be discussed.

First, it is possible that the innovative speakers might simply – just because they more advanced than their peers in the change – have lost to some extent the capacity to voice /v/ over time and therefore fail to imitate those in the experiment, even if they are instructed to. Voiced fricatives are relatively rare cross-linguistically (Ohala 1983) and they are notably challenging – especially in initial position – for the articulatory system as they require at the same time a pressure drop across the constriction to generate noise and a maintenance of pressure across the vocal folds in order to generate voicing (e.g., Stevens 1971; Ohala 1997). This is probably the key phonetic pre-condition for this sound change to initiate in the first place. Dutch speakers however produce voiced /v/ in other linguistic contexts in a very stable manner (f.i., intervocalic position), thus a simple 'articulatory' loss of the ability to produce voicing cannot be attested. The extent to which phonological environments constraint the ability to produce a sound constitutes an interesting further step in this line of research.

Besides voicing which was the central cue in this study, we reviewed the role of additional phonetic cues in the labiodental fricative contrast /f/-/v/. Like in previous work (e.g., Kissine et al. 2003, 2005; Pinget 2015), this study showed that voiced fricatives become as long as their voiceless counterparts as a result of the change. F0 contours were shown to be typically higher after /f/ than after /v/. This effect was also found in the forced imitation task (where the cue was neutralized in the input), which is giving indirect, but strong evidence for an automatic and articulatory-motivated account of F0 perturbations after obstruents (as proposed by e.g., Halle and Stevens 1971; Hanson 2009). Limiting the discussion to the main phonetic cue of the amount of voicing under consideration here, we observed that there was no significant interaction between the main effect of speaker's innovativeness with the factor region. Such an interaction would however be expected if speakers' advancement in the change was the cause of the diminished imitation capacity, as speakers from Groningen as a group are significantly further in the change than speakers from West-Flanders.

All in all, the second option seems more favorable: it is exactly this diminished accuracy in phonetic imitation that allows some individual speakers to be more advanced in the change. These speakers are generally worse at reproducing the existing speech patterns in a precise way, and therefore make the sound change in their own region possible. The other speakers, in contrast, show a large phonetic accuracy which allows them to precisely reproduce the speech patterns in their region, and therefore tend to reinforce the stability of phonetic realizations within the community. We favor this explanation as it seems to offer a direct explanation for the fact that linguistic systems are both stable and changing: good 'forced' imitators in general are responsible for keeping stability in the system, while poor 'forced' imitators make sound change possible, possibly because they are capable of introducing new patterns into the system. Consensuses are made again and again within each conversation between these two tendencies. An interesting way to test the direction of causation here would be to test forced imitation capacities in different phonetic contrasts in the same speakers. This could offer insight into the extent to which this imitation capacity is a general linguistic characteristic of speakers (and therefore not related to a specific fricative contrast only) and can be used to either stabilize or enhance any (other) sound change. Tamminga (2021) already asked the question whether the same individuals might lead different changes and has shown that investigating interspeaker covariation patterns might provide some answers to that question. The current study put forward the importance of investigating imitation capacities additionally to or along with those covariation patterns. Tamminga (2021) also discussed the possible discrepancy

between studies in which speaker-specific degree of innovativeness throughout a community is quantified and traditional sociolinguistic studies in which linguistic leaders (saccadic leaders in Labov's terminology (2001:384)) are portrayed. Along with her study, the current work underlines the needs of quantifying the distribution of speaker differences across the community level, and of bridging the gap between community-based sociolinguistic research and laboratory phonology imitation studies.

The spontaneous imitation task (Experiment 3) concentrated on participants' readiness to deviate in social interaction. The task demonstrated – like many previous imitation studies (e.g., Babel 2012; Delvaux and Soquet 2007; Dufour and Nguyen 2013) – that simple exposure to model speakers can induce spontaneous imitation. Only a few speakers however showed throughout the whole experiment a significant tendency to converge, which supports the findings of Sonderegger et al. (2017). The induced modifications at the item level, measured as the distance between the imitated and the baseline productions, showed large individual differences which correlate with the position of speakers in the change advancement. Leaders of change in each region turned out to be more inclined to imitate the model talkers and to produce the new variant than the more conservative speakers. Again, because this effect was stable in all three regions, we have serious reasons to think that it is exactly this enhanced readiness to go along with the change that explains why leaders are further in the change. They were more inclined to find consensuses with other speakers (even the ones showing a deviant/new form) on the phonetic level. In contrast, the more conservative speakers were less inclined to spontaneously converge towards the new variant and stuck to their own variant, sustaining the old system.

Both imitation tasks were complementary in order to gain insight into both the articulatory and auditory constraints on the imitation process and the more social aspects of imitation. In the forced imitation task, we might safely assume that the results reflect the highest level of imitation speakers might achieve, as they were clearly instructed to do so. Even if it was designed to measure participants' capacity to produce voicing, it cannot be ruled out that participants still differed in the extent to which they choose to do so or in the 'effort' they made to follow the instructions, besides the question whether or not they were capable to imitate. In contrast, the extent to which speakers have spontaneously imitated the model talkers in Experiment 3 was supposed to be unconscious (or at least not instructed and not explicit) and possibly modulated by a large range of factors like the attractiveness of model talkers, their gender, etc (see literature review in Section 1.3.). Kim et al. (2011) also reported more spontaneous imitation when

speakers and model talkers share the same dialect backgrounds. Our model talkers shared background with some speakers, but with not all. This might possibly have resulted in individual differences in the spontaneous imitation task. There is however no way to control for this influence in a design where speakers from different regions are included. In addition to these issues, there was also a substantial difference between the forced and spontaneous imitation tasks in the nature of the stimuli (manipulated voicing continuum vs. naturally produced devoiced fricatives). Future studies might need to tackle these issues in order to better compare different types of imitation capacities. Note also that participants in this study were labelled as ‘leaders of change’, but this qualification is only a relative one. They are ‘leaders’ compared with the other speakers in the study and not necessarily leaders for the whole community, as we do not know the exact extent to which the regional patterns in the study are representative of that region. Future work could show how advanced speakers are in a sound change, judged against the immediate community with which they interact (for instance based on the immediate networks of friends and family). However, it is a challenge to implement this kind of procedure into a feasible research design.

The topic of speech perception was not directly mentioned in this study up onto this point. However, it is clear that there cannot be imitation if the variant participants are asked or supposed to imitation is not first perceived. The role of speech perception lies thus at the core of this investigation, just as it lies at the core of much research on sound change, ranging from Ohala’s model of sound change initiation where unperfect perception is responsible for change (Ohala 1983) onto Beddor (2012) showing individual differences in the way subjects align their perceptual and production systems (see an overview of the literature by Stevens and Harrington 2014). Testing phonetic imitation is also indirectly assuming some link between speech perception and production systems. Research about the nature of this link, its strength and the presence of individual differences in how speech perception align to speech production goes hand in hand with the current study. Importantly, we need to define the role of the imitation process in the systems. Goldinger (1997) for instance proposed that – upon hearing a word – all its episodic traces are activated in memory and that this activation creates a ‘generic echo’ or the mean of the activated set. It is the mean of the activated set that is selected for production. So, phonetic imitation occurs when auditory exposure to a model talker causes the speaker’s productions to shift towards those of the model talker (Goldinger 1997).

However, the existence of individual differences in imitation cannot directly be explained when imitation takes such an intrinsic position in the structure of the linguistic system. Additional insights into speakers' more general cognitive abilities might be necessary in order to achieve a good explanation for those individual differences. Possible new paths include measuring imitation of speech patterns that are not related to the sound change under consideration, or imitation of more general aspects of speech (e.g., speech rate, loudness, etc.). In general, future work might profit from attempts to disentangle between imitation as the core mechanism relating speech production with speech perception and imitation as a more general cognitive ability showing individual differences.

In conclusion, this paper highlighted individual differences in imitation capacities, which were correlated to the spread of a sound change within communities. Leaders of sound change were shown to be less accurate in forced phonetic imitation than conservative speakers, but more inclined to spontaneously imitate model talkers. These insights exemplify a recent line of sociophonetic research testing the idea that variability between individuals may be the key to understanding how sound change comes about and trying to relate the two apparently paradoxical features of the language system: its stability and its potential for sound change.

Acknowledgments: I 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. I am also very grateful to Hans Rutger Bosker, Carolien van Hazelkamp and Elise Drijbooms whose voices were 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. Thanks to René Kager, Hans Van de Velde, Willemijn Heeren, Hanna Ruch and Cesko Voeten for their comments on drafts of this paper. Finally, we are very grateful to the editors and the anonymous reviewers for their feedback on earlier versions of their manuscript. All remaining errors and shortcomings are our own.

Research funding: This work was supported by the Netherlands Organization for Scientific Research (NWO grant 322-75-002).

Conflict of Interest Statement: The author has no conflicts of interest to declare.

Appendix 1

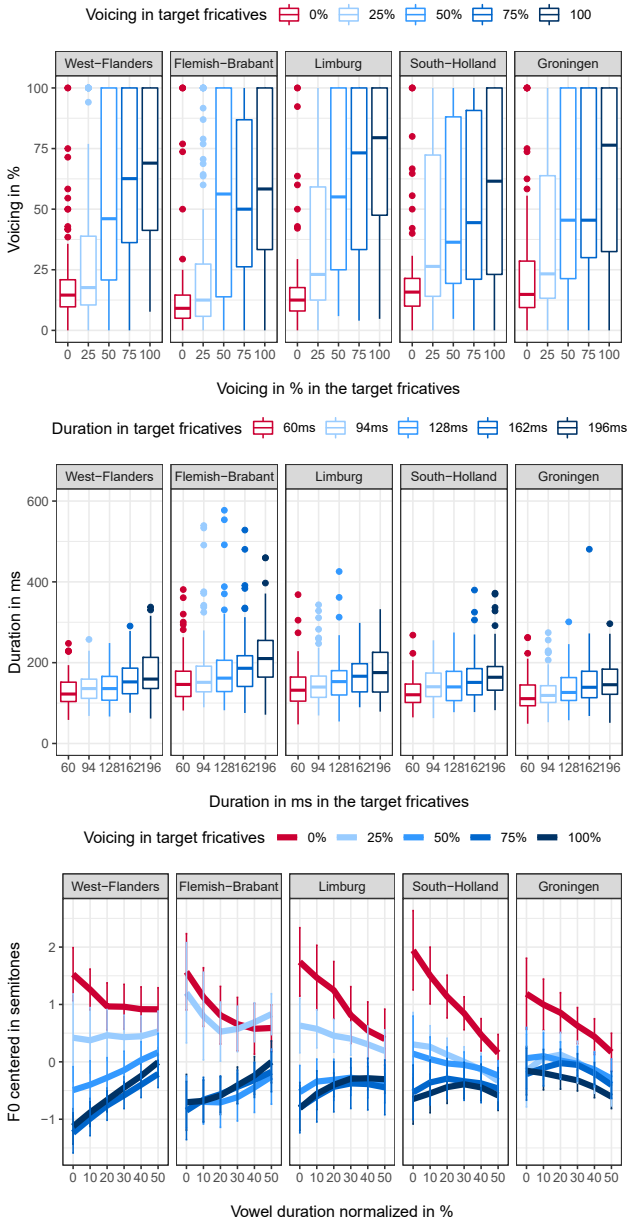


Figure A: Three measured phonetic cues in the forced imitation experiment. Top: Boxplot of voicing measures (in %) depending on the amount of voicing in the target fricatives, split up by region. Middle: Boxplot of duration measures (in ms) depending on the duration in the target fricatives, split up by region. Bottom: Line graphs of F0 centered, in semitones, in the first 50% of the following vowel, split up by region (bars represent the confidence interval).

Appendix 2

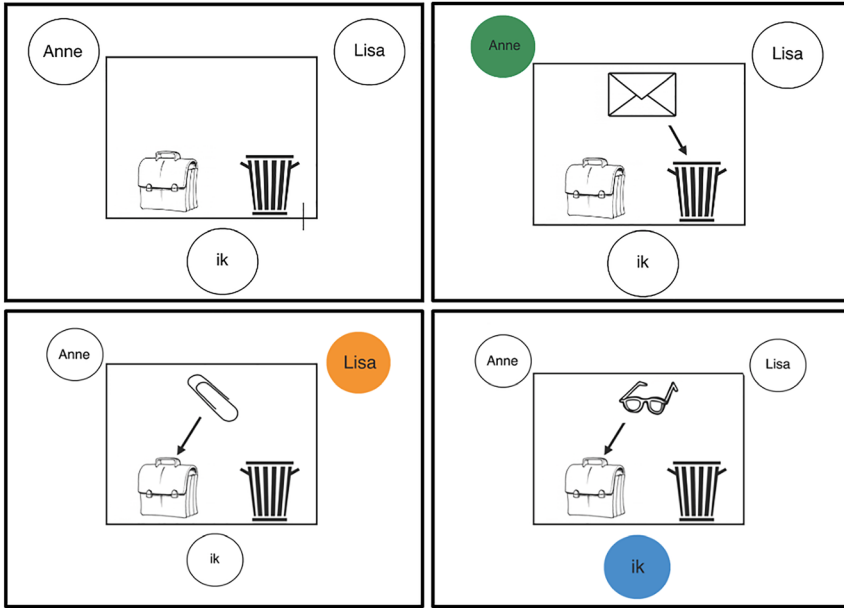


Figure B: Screen shots of the spontaneous imitation task: four trials (top left: baseline screen, top right: trial A, bottom left: trial L and bottom right: trial P).

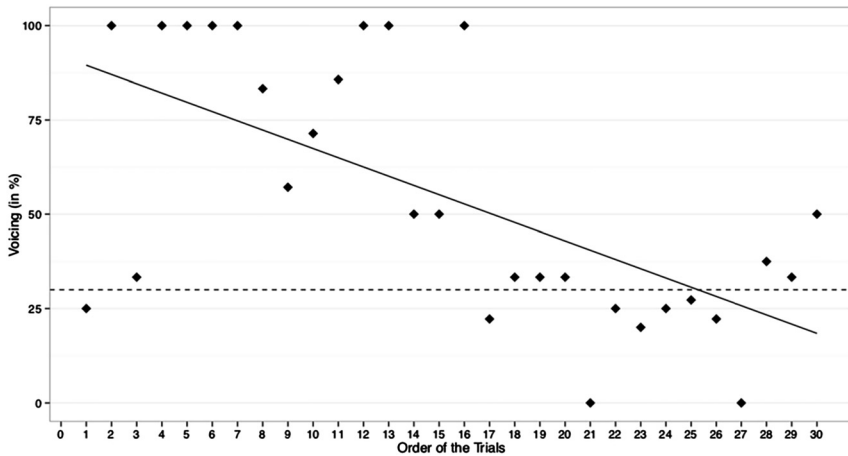


Figure C: Example of phonetic convergence in the spontaneous imitation task. Fricative realizations of participant #LI20 in the spontaneous imitation task. The order of trials is presented along the x-axis and the amount of voicing along the y-axis. The dotted line represents the average voicing produced by the model talkers and the full line is the fitted regression line.

References

- Abel, Jennifer & Molly Babel. 2017. Cognitive load reduces perceived linguistic convergence between dyads. *Language and Speech* 60(3). 479–502.
- Abrego-Collier, Carissa, Julian Grove, Morgan Sonderegger & C. L. Alan. 2011. Effects of speaker evaluation on phonetic convergence. In *ICPhS*, 192–195.
- Ash, Sharon. 2002. Social class. In J. K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*, 402–422. Oxford: Blackwell.
- Auer, Peter & Frans Hinskens. 2005. The role of interpersonal accommodation in a theory of language change. In Peter Auer, Frans Hinskens & Kerswill Paul (eds.), *Dialect change: Convergence and divergence in European languages*, 335–357. Cambridge, UK: Cambridge University Press.
- Babel, Molly. 2010. Dialect divergence and convergence in New Zealand English. *Language in Society* 39(4). 437–456.
- Babel, Molly. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 40(1). 177–189.
- Baker, Adam, Diana Archangeli & Jeff Mielke. 2011. Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change* 23. 347–374.
- Bates, Douglas, Martin Maechler, Bolker Ben & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67. 1–48.
- Beddor, Patrice Speeter. 2012. Perception grammars and sound change. In Maria-Josep Solé & Daniel Recasens (eds.), *The initiation of sound change: Production, perception, and social factors*, 37–55. Amsterdam, Netherlands: John Benjamins.
- Boersma, Paul & David Weenink. 2022. *Praat: Doing phonetics by computer* [Computer program]. Version 6.2.23. <http://www.praat.org/> (accessed 8 October 2022).
- Cassier, Luc & Pierre Van de Craen. 1986. Vijftig jaar evolutie van het Nederlands [Fifty years evolution of the Dutch language]. In Jos Cretien, Geerts Guido & Koen Jaspaert (eds.), *Momentopnamen van de sociolinguïstiek in België en Nederland*, 59–73. Leuven, Belgium: Acco.
- Chambers, J.K. 2002. Patterns of variation including change. *The handbook of language variation and change*. UK: Blackwell Publishing Ltd.
- Cohen, Antonie, Carl L. Ebeling, Klaas Fokkema & André van Holk. 1961. *Fonologie van het Nederlands en het Fries* [Phonology of Dutch and Frisian]. Den Haag, the Netherlands: Martinus Nijhoff.
- Croft, William. 2000. *Explaining language change: An evolutionary approach*. London: Longman.
- Delvaux, Véronique & Alain Soquet. 2007. The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica* 64(2–3). 145–173.
- Drager, Katie & Jennifer Hay. 2012. Exploiting random intercepts: Two case studies in sociophonetics. *Language Variation and Change* 24(1). 59–78.
- Dufour, Sophie & Noël Nguyen. 2013. How much imitation is there in a shadowing task? *Frontiers in Psychology* 4. 346.
- Eckert, Penelope & Sally McConnell-Ginet. 2003. Language and gender. *Language* 80(4). 846–849.
- Garrett, Andrew & Keith Johnson. 2013. Phonetic bias in sound change. In Alan Yu (ed.), *Origins of sound change: Approaches to phonologization*, 51–97. Oxford: Oxford University Press.

- Giles, Howard & Philip Smith. 1979. Accommodation theory: Optimal levels of convergence. In Howard Giles & Robert St Clair (eds.), *Language and social Psychology*, 45–65. Baltimore: Basil Blackwell.
- Giles, Howard, Donald Taylor & Richard Bourhis. 1973. Towards a theory of interpersonal accommodation through language: Some Canadian data. *Language in Society* 2. 177–192.
- 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.
- Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105(2). 251.
- Gordon, Matthew, Barthmaier Paul & Kathy Sands. 2002. A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association* 32. 141–174.
- Gussenhoven, Carlos. 1999. Illustrations of the IPA: Dutch. In *Handbook of the international phonetic association*, 74–77. Cambridge: Cambridge University Press.
- Halle, Morris & Kenneth N. Stevens. 1971. A note on laryngeal features. *Quarterly Progress Report of the Research Laboratory of Electronics*, M.I.T 101. 198–213.
- Hamann, Silke & Anke Sennema. 2005. Acoustic differences between German and Dutch labiodentals. *ZAS Papers in Linguistics* 42. 33–41.
- Hanson, Helen M. 2009. Effects of obstruent consonants on fundamental frequency at vowel onset in English. *Journal of the Acoustical Society of America* 125(1). 425–441.
- Harrington, Jonathan, Sallyanne Palethorpe & Catherine I. Watson. 2000. Does the queen speak the queen's English? *Nature* 408(6815). 927–928.
- Harrington, Jonathan & Florian Schiel. 2017. /u/-fronting and agent-based modeling: The relationship between the origin and spread of sound change. *Language* 93(2). 414–445.
- Jassem, Wiktor. 1979. Classification of fricative spectra using statistical discriminant functions. In Björn Lindblom & Sven Öhman (eds.), *Frontiers of speech communication research*, 77–91. New York: Academic Press.
- Kim, Midam, William S. Horton & Ann R. Bradlow. 2011. Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology* 2(1). 125–156.
- Kissine, Mikhail, Hans Van de Velde & Roeland van Hout. 2003. An acoustic study of standard Dutch /v/, /f/, /z/ and /s/. *Linguistics in the Netherlands* 20(1). 93–104.
- Kissine, Mikhail, Hans Van de Velde & Roeland van Hout. 2005. Acoustic contributions to sociolinguistics: Devoicing of /v/ and /z/ in Dutch. *University of Pennsylvania Working Papers in Linguistics* 10(2). Article 12.
- Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2(2). 205–254.
- Labov, William. 1994. *Principles of language change: Internal factors*, 1. Oxford: Blackwell.
- Labov, William. 2001. *Principles of linguistic change: Social factors*, 2. Oxford: Blackwell.
- Lev-Ari, Shiri. 2018. Social network size can influence linguistic malleability and the propagation of linguistic change. *Cognition* 176. 31–39.
- Mees, Inger & Beverly Collins. 1982. A phonetic description of the consonant system of standard Dutch. *Journal of the International Phonetic Association* 12. 2–12.
- Mitterer, Holger. 2009. Research stuff. <http://www.holgermitterer.eu/research.html> (accessed 17 May 2015).
- Milroy, James & Lesley Milroy. 1985. Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21(2). 229–284.

- Milroy, Lesley. 1987. *Language and social networks*. Oxford: Blackwell.
- Milroy, Lesley. 2002. Introduction: Mobility, contact, and language change – Working with contemporary speech communities. *Journal of Sociolinguistics* 6(1). 3–15.
- Namy, Laura L., Lynne C. Nygaard & Denise Sauerteig. 2002. Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology* 21(4). 422–432.
- Niedzielski, Nancy & Howard Giles. 1996. Linguistic accommodation. In Hans Goebel, Peter H. Nelde, Zdenek Stary & Wolfgang Wölck (eds.), *Kontaktlinguistik: Ein internationales Handbuch zeitgenössischer Forschung* [An International Handbook of Contemporary Research], vol. 1, 332–342. Berlin: Mouton de Gruyter.
- Nielsen, Kuniko. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39(2). 132–142.
- Ohala, John. 1997. Aerodynamics of phonology. *Proceedings of the Seoul International Conference on Linguistics* 92. 97.
- Ohala, John. 1983. The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer-Verlag.
- Pardo, Jennifer S. 2006. On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America* 119(4). 2382–2393.
- Pardo, Jennifer S. 2013. Measuring phonetic convergence in speech production. *Frontiers in Psychology* 4. 559.
- Pardo, Jennifer S., Adelya Urmanche, Sheryl Wilman, Jaclyn Wiener, Nicholas Mason, Kaegan Francis & Melanie Ward. 2018. A comparison of phonetic convergence in conversational interaction and speech shadowing. *Journal of Phonetics* 69. 1–11.
- Paul, Hermann. 1880. *Prinzipien der Sprachgeschichte* [Principles of language history]. Niemeyer: Tübingen.
- Pinget, Anne-France. 2015. *The actuation of sound change*. PhD Dissertation Utrecht University. LOT series.
- Pinget, Anne-France, René Kager & Hans Van de Velde. 2020. Linking variation in perception and production in sound change: Evidence from Dutch obstruent devoicing. *Language and Speech* 63(3). 660–685.
- Pinget, Anne-France & Hugo Quené. 2021. Effects of obstruent voicing on vowel fundamental frequency in Dutch. In *Paper presented at the Phonetics day*. The Netherlands: Dutch Association for Phonetic Sciences.
- Sankoff, Gillian & Hélène Blondeau. 2007. language change across the lifespan: /r/ in Montreal French. *Language*. 560–588. <https://doi.org/10.1353/lan.2007.0106>.
- Shultz, Amanda A., Alexander L. Francis & Fernando Llanos. 2012. Differential cue weighting in perception and production of consonant voicing. *Journal of the Acoustical Society of America* 132. EL95–EL101.
- Sievers, Eduard. 1901. *Grundzüge der Phonetik zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen* [Foundations of phonetics as an introduction to the study of the phonetics of the Indo-European languages]. Leipzig: Breitkopf & Härtel.
- Slis, Iman H. & Antonie Cohen. 1969. On the complex regulating the voiced-voiceless distinction I and II. *Language and Speech* 12. 80–102.
- Slis, Iman H. & Marieke van Heugten. 1989. Voiced-voiceless distinction in Dutch fricatives. In Hans Bennis & Ana van Kemenade (eds.), *Linguistics in The Netherlands*, vol. 6, 123–132.
- Sonderegger, Morgan, Max Bane & Peter Graff. 2017. The medium-term dynamics of accents on reality television. *Language* 93(3). 598–640.

- Stevens, Kenneth N. 1971. Airflow and turbulence noise for fricative and stop consonants: Static considerations. *Journal of the Acoustical Society of America* 50(4). 1180–1192.
- Stevens, Mary & Jonathan Harrington. 2014. The individual and the actuation of sound change. *Loquens* 1(1). e003.
- Tammenga, Meredith. 2021. Leaders of language change: Macro and micro perspectives. In Hans Van de Velde, Nanna H. Hilton & Remco Knooihuizen (eds.), *Language variation – European perspectives VIII: Selected papers from the tenth international conference on language variation in Europe (ICLaVE 10)*, 270–289. Amsterdam: John Benjamins Publishing Company.
- Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Blackwell.
- Trudgill, Peter. 2004. *New-dialect formation: The inevitability of colonial Englishes*. Edinburgh: Edinburgh University Press.
- Trudgill, Peter. 2008. Colonial dialect contact in the history of European languages: On the irrelevance of identity to new-dialect formation. *Language in Society* 37(2). 241–254.
- Van de Velde, Hans. 1996. *Variatie en verandering in het gesproken Standaardnederlands [Variation and change in Spoken Standard Dutch] (1935–1993)*. University of Nijmegen PhD Dissertation.
- 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 der Wal, Marijke & Cor van Bree. 1992. *Geschiedenis van het Nederlands [History of the Dutch language]*. Utrecht: Spectrum.
- van Son, Rob. 2000. Protocol voor het oplijnen van fonetische transcripties met spraak. Available at: <http://www.fon.hum.uva.nl/IFA-SpokenLanguageCorpora/IFACorpus/SLcorpus/LabelProtocol/LabelProtocol.pdf>.
- van Son, Rob & Louis Pols. 1996. An acoustic profile of consonant reduction. *Proceedings of the Fourth International Conference on Spoken Language* 3. 1529–1532.
- Voeten, Cesko C. 2021. Individual differences in the adoption of sound change. *Language and Speech* 64(3). 705–741.
- Weinreich, Uriel, William Labov & Marvin Herzog. 1968. *Empirical foundations for a theory of language change*. Austin: University of Texas Press.
- Yu, Alan. 2013. Individual differences in socio-cognitive processing and sound change. In Alan Yu (ed.), *Origins of sound change: Approaches to phonologization*, 201–227. Oxford University Press.
- Yu, Alan, Carissa Abrego-Collier & Sonderegger Morgan. 2013. Phonetic imitation from an individual-difference perspective: Subjective attitude, personality and “autistic” traits. *PLoS One* 8(9). e74746.