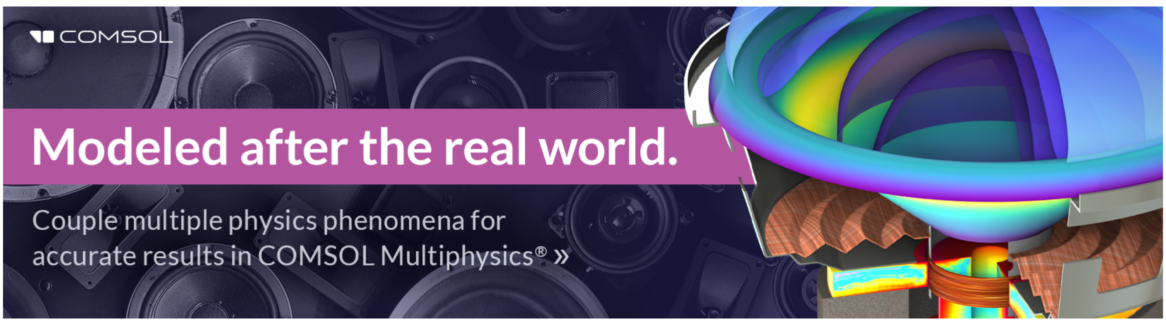
The Journal of the Acoustical Society of America

OCTOBER 05 2023

Effects of obstruent voicing on vowel fundamental frequency in Dutch

Anne-France Pinget; Hugo Quené

J. Acoust. Soc. Am. 154, 2124–2136 (2023)
https://doi.org/10.1121/10.0021070



JASA ARTICLE

Effects of obstruent voicing on vowel fundamental frequency in Dutch

Anne-France Pinget[a)] and Hugo Quené
Institute of Language Sciences, Utrecht University, Trans 10, 3512 JK Utrecht, the Netherlands

ABSTRACT:
It has been known for a long time and a wide variety of languages that vowel fundamental frequency (F0) following voiceless obstruents tends to be significantly higher than F0 following voiced obstruents. There has been a long-standing debate about the cause of this phenomenon. Some evidence in previous work is more compatible with an articulatory account of this effect, while others support the auditory enhancement account. This paper investigates these consonant-related F0 perturbations in Dutch after initial fricatives (/v, f/) and stops (/b, p/), as compared to after the nasal /m/. Dutch is particularly interesting because it is a "true voicing" language, and because fricatives are currently undergoing a process of devoicing. Results show that F0 was raised after voiceless, but largely unaffected after voiced obstruents. Fricative voicing in /v/ and F0 level tend to covary: the less voicing in /v/, the higher F0 at onset. There was no trace of an active gesture to explicitly lower pitch after highly devoiced fricatives, as would be predicted by an auditory account. In conclusion, F0 perturbations after Dutch obstruents and their covariation patterns are taken as additional evidence to support an articulatory cause of consonant-related F0 effects.
 https://doi.org/10.1121/10.0021070

(Received 11 April 2023; revised 28 August 2023; accepted 29 August 2023; published online 5 October 2023)

[Editor: Susanne Fuchs] Pages: 2124–2136

I. INTRODUCTION

In a wide variety of languages, it has been well documented that fundamental frequency (F0) at the onset of a vowel following a voiceless obstruent is higher than at the onset of a vowel following a voiced obstruent (for English see House and Fairbanks, 1953; Kingston and Diehl, 1994; Lehiste and Peterson, 1961; Mohr, 1971, and in many other languages like Korean, Swedish, French, Italian, Danish, Xhosa; see Ting *et al.*, 2023 for a recent review and comparison of the literature). This effect has received a variety of different names, sometimes called "pitch skip" (Hanson, 2009) or "consonants-related F0 perturbations" (Kirby and Ladd, 2016), "onset voicing effect" (Kirby and Ladd, 2015), or "consonant intrinsic F0" (Kingston, 2007). Following Krug *et al.* (2021), we refer to this effect as *co-intrinsic F0* (CF0). While CF0 effects have meanwhile been studied extensively, there is a long-standing debate regarding their underlying source.

This paper investigates CF0 effects in Dutch obstruents. Dutch stops are of interest because the contrast lies between voiceless unaspirated stops and fully voiced (prevoiced) stops, i.e., Dutch is a *true voicing* language. F0 after Dutch fricatives has—to our current knowledge—not been the subject of a thorough investigation since the study of Slis and Cohen (1969). Yet, Dutch fricatives are highly relevant to our understanding of CF0 effects, as they are currently involved in a well-document process of devoicing where voiced fricatives are produced as (partially) voiceless. This paper presents a detailed investigation of F0 contours, the magnitude and time course of CF0 effects, and the covariation patterns between phonetic cues in Dutch obstruents, aiming to contribute to the debate about the source of CF0 effects.

A. Source of CF0 effects: A long-standing debate

Despite the large number of studies showing that voicing contrasts in many languages are accompanied by effects on the F0 of the following vowel, there is still debate among researchers about the source of these effects. The proposed causes for F0 differences fall into two clusters: (1) auditory-motivated accounts and (2) articulatory-motivated accounts.

On the one hand, Kingston and Diehl (1994) proposed an auditory-motivated account to explain F0 differences. They suggested that speakers actively lower F0 following a voiced obstruent to signal the [+voice] feature to listeners. Because it is based on the principle of contrast enhancement, this explanation can be considered of a phonological nature. This view is also supported by Ohde (1984) who reported that in English a high F0 at onset was found for both voiceless aspirated and unaspirated plosives, although the VOTs were very different. They suggested that F0 differences are a product of articulations that are controlled independent of the timing of the glottal articulations to produce voicing. Hence this effect is the consequence of a controlled, deliberate enhancement of [+voice]. We will refer to this account as the auditory enhancement account as it is related to how listeners integrate acoustic cues perceptually.

[a)]Also at: Fryske Akademy, Doelestraat 8, 8911 DX Leeuwarden, the Netherlands.

On the other hand, there is reason to believe that this effect may simply fall out from the physiology or articulation of obstruent production. F0 differences might directly reflect the automatic effect of laryngeal articulatory gestures and are therefore articulatory-motivated (Hanson 2009). Halle and Stevens (1971) specifically proposed that voiceless obstruents tend to raise F0 of the following vowel because their production canonically involves a stiffening of the vocal folds. This stiffening primarily results in the inhibition of phonation, but it also tends to increase vocal fold length which may result in a rise in F0 at the onset of the following segment. This view in which the F0 perturbations are caused by the physiology of voiceless obstruents also received support from Löfqvist *et al.* (1989), who found that cricothyroid muscle activity increases for voiceless consonants relative to voiced consonants. An increase in cricothyroid activity intensifies the longitudinal tension of the folds, which helps to inhibit voicing during the voiceless consonant and also increases the frequency of vibration upon phonation for the following vowel. Hombert *et al.* (1979), however, showed that there is possibly another force at play: F0 is lowered after voiced obstruents. They proposed that a possible cause for the F0 perturbations is the lowering of the larynx during the production of voiced plosives in order to obtain sufficient transglottal pressure to produce vocal fold vibration. Honda *et al.* (1999) gave additional support for this hypothesis, showing that the lowering of the larynx can cause a downward tilt of the cricoid cartilage, which causes a shortening of the vocal folds and thereby a decrease in their horizontal tension. These gestures result in a lower F0 in voiced obstruents that carries over into the adjacent vowel. We will refer to this account as the articulatory-motivated account.

Several later studies providing general evidence for an articulatory account of CF0 differences further tried to disentangle these two possible physiological forces (F0 raised after voiceless obstruents vs lowered after voiced obstruents) by introducing a nasal baseline condition (Hanson, 2009; Kirby and Ladd, 2016). It is generally accepted that the production of nasals does not require any articulatory adjustments of the supralaryngeal cavity and that the airflow going through the nasal cavity does not perturb F0 (Ohala, 1975; Hombert *et al.*, 1979). Hanson's (2009) results on English using a nasal /m/ as a reference level pointed towards a raised F0 after voiceless consonants, as the F0 contour was higher than F0 after the nasals long into the vowel. Looking at true voicing languages Italian and French, Kirby and Ladd (2016) provided evidence for both physiological forces: they showed both raised F0 following the release of voiceless consonants and depressed F0 during the closure phase of voiced obstruent, as compared to the nasal reference level. Working along the same lines as Hanson (2009) and Kirby and Ladd (2016), the current study includes a nasal baseline in order to investigate—in the case of an articulatory-motivated explanation for CF0 effects—which physiological force is at play, i.e., to establish the change in F0 after Dutch obstruents relative to F0 after Dutch sonorants, which we regard as an unperturbed reference level.

B. Covariation patterns between cues

The two above-described accounts make different predictions about how phonetic cues covary and provide different explanations about why they do so. For stops, the articulatory account predicts a correlation between voice onset time (VOT) and onset F0 in the following vowel that automatically follows from the articulatory settings involved in voicing production, and that is not directly controlled by the speaker (Hombert *et al.*, 1979; Löfqvist *et al.*, 1989). In contrast, the auditory enhancement account predicts that that the connection between these two cues is intentional and phonologically-determined (Keating, 1984; Kingston and Diehl, 1994; Kingston, 2007). The F0 cue helps to enhance the perception of segments as [+voice] stops, thereby increasing the perceptual distinctiveness between [+voice] and [−voice] stops.

Covariation patterns between VOT and onset F0 have been examined for stops in different types of languages, but the results are contradictory. Shultz *et al.* (2012), for instance, presented data from a production study of 32 native speakers of American English and showed a significant inverse correlation between VOT and onset F0, suggesting that the F0 effect was attenuated by speakers who produced stops with longer VOTs. The authors interpreted their findings as being consistent with the account from Kingston and Diehl (1994). Like Repp (1982), they suggested that VOT and onset F0 are in a trading relationship: a change in the value of one cue that would otherwise result in a different phonetic percept can be offset by a change in the value of another. Yet, the Kirby and Ladd (2015) data on French and Italian showed that rather than an inverse correlation between voicing lead and onset F0, longer prevoicing is accompanied by a lower F0 at the onset of the following vowel (based on words in isolation, thus utterance-initial stops). They interpreted these findings as more compatible with the articulatory account: this relationship is expected if F0 effects are a by-product of laryngeal lowering, a gesture maintaining the transglottal pressure to maintain voicing during the stop closure. This finding was, however, not replicated in Kirby and Ladd (2016), a follow up study on the same languages, but involving different elicitation context, i.e., carrier phrases. Kirby and Ladd (2016) also reported considerable individual variation in the strength and direction of correlations, a point also made by Dmitrieva *et al.* (2015) and Gao and Arai (2019). Dmitrieva *et al.* (2015) moreover reported within-category VOT and onset F0 to be uncorrelated in Spanish and in English data. In conclusion, the examination of how VOT and onset F0 co-vary in different languages has yielded conflicting results.

In general, the study of CF0 effects has largely concentrated on stops, at the expense of our knowledge about other obstruents. This is also true regarding covariation patterns. Kirby and Ladd (2016) did look at the correlations between noise duration and F0 at vowel onset for fricatives and found no significant correlation. This study therefore includes—in addition to the covariation patterns for Dutch stops—an examination of covariation patterns for fricatives, and thus aims to look for additional support for either the articulatory or the auditory enhancement account (see hypotheses under Sec. I F).

C. Characteristics of the CF0 effects: Time course, contours, and magnitude

A possible reason why there is still a debate about the source of CF0 effects is the lack of clarity about how to characterize these effects (Ladd and Schmid, 2018). Studies tend to differ as to which aspects they focus on: either the time course of CF0 effects, the contour shapes, or the magnitude of CF0 differences.

Regarding the time course of CF0 effects, House and Fairbanks (1953) started with measurements of the average F0 over the whole vowel segment, but already noted that F0 differences between voiced and voiceless consonants did not occur uniformly through the following vowel but appeared mostly at the onset of the vowel. Mohr (1971) measured maximum vs minimum F0 over the entire duration of vowels and suggested that the influence of the obstruents is limited to the early part of the vowel. However, looking at languages other than English, it was shown that F0 perturbations might extent further into the vowel. Hombert (1978) reported perturbations for Yoruba that extend up to 100 ms into the vowel. Kenstowicz and Park (2006) looked at Kyungsang Korean reported perturbations at least until mid-vowel. In contrast, Francis *et al.* (2006) reported that F0 perturbations in Cantonese—more like their English counterparts—are more limited to the very early portion of the vowel (first 0–10 ms). As these studies also significantly differ in terms of type of recordings, these findings are difficult to compare. Hence, there is no consensus whether the obstruent effect only involves the absolute F0 value immediately after the consonants or whether it entails an F0 movement into the vowel, and if so, until which part of the vowel. Yet, the question about the time course of CF0 perturbations is crucial to understanding which mechanisms are at force and what the source of the effect is. We will return to this point in the discussion.

Many researchers have focused more specifically on the contour shapes of CF0 effects, noting that F0 falls after voiceless obstruents and rises after voiced ones (e.g., Whalen *et al.*, 1999). Hombert *et al.* (1979) also focused on the contour shapes showing that F0 contours do not slope down for all speakers, making an important point regarding individual differences. Some individual variability in the F0 contours was also shown by Löfqvist *et al.* (1995). It is very plausible that the shape of CF0 differences is the most perceptually relevant feature and that should be a crucial argument for the auditory enhancement account. Haggard *et al.* (1970), for instance, showed that stops with a low-rising F0 contour in the following vowel are consistently perceived as /b/, but as /p/ with a high-falling contour in the vowel, underlying the perceptual importance of F0 movement into the vowel as well. However, other perceptual evidence provided by Haggard *et al.* (1981) suggested that the actual perceptual cue listeners attend to is the onset frequency, rather than the F0 movement into the vowel. Hanson (2009) investigated CF0 effects in production after English obstruents and measured both the onset of the vowel and the F0 contours throughout the vowel. She claimed that the effects are indeed not so much related to the F0 contours, but mostly related to the absolute F0 difference over the first few tens of milliseconds in the vowel. The magnitude of CF0 differences thus appears to be crucial. However, this aspect is the most challenging as it is not straightforwardly comparable across studies, since we know that it highly depends on many factors, such as the general intonational context, the vowels under consideration, the gender of the speakers, etc. (Chen, 2011; Hanson, 2009; Kirby, 2018; Kirby and Ladd, 2016; Gao and Arai, 2019).

To summarize, previous studies tend to focus on a certain aspect of CF0 effects instead of providing a comprehensive analysis (except for the recent study by Ting *et al.*, 2023, which includes several aspects in 16 different languages). Moreover, heterogeneity in measurements makes the comparison between studies and the evaluation of their contribution to the theoretical debate challenging. The current study aims to examine all aspects of CF0 effects (i.e., the magnitude of the CF0 differences, F0 contours onto the entire vowel, and thus the time course of the effects). It will include a large sample of speakers for whom individual differences are taken into account. Generalized Additive Mixed Modeling (GAMM) (Wood, 2017; Wieling, 2018) will be used to assess and compare smoothed F0 contours in various CF0 contexts, including a baseline context, while allowing for individual differences between speakers in their F0 contours.

D. Dutch as a true voicing language

While CF0 effects are thought to be found cross-linguistically and regardless of how the voicing contrasts are mapped onto phonetic differences in the VOT distribution (Ting *et al.*, 2023), most of the CF0 evidence is based on aspirating languages, very often English. English traditionally contrasts between word-initial short lag voiced stop and long lag voiceless stop (Lisker and Abramson, 1964; Cho and Ladefoged, 1999). Kirby and Ladd (2015, 2016) underlined the importance of investigating other languages, especially true voicing languages (also called true voice languages by Beckman *et al.*, 2013). They suspected that F0 lowering may be more easily observed when there is vigorous voicing during the closure, a characteristic that is often called true voicing. In English, the [voice] feature in stops primarily manifests itself through VOT. Voicing during closure in stops generally does not affect the phonological status of the stop, although this may depend on the context and the specific variety of English (Davidson, 2016; Flege 1982; Jacewicz *et al.*, 2009). In recent years, some criticism against the strictly binary interpretation of VOT has been expressed (e.g., Ladd and Schmid, 2018, for a critique of "laryngeal realism" based on data from Swiss German, where closure duration, not VOT is the primary cue to the fortis vs lenis contrast between stops). We will return to this topic in the discussion.

Dutch in contrast is relatively unusual within the Germanic languages in that it has a true voicing contrast (Keating, 1984; Lisker and Abramson 1964).[1] Along with other languages such as Arabic, Japanese, Polish, French, Italian, Spanish, Russian or Hungarian, the contrast in standard Dutch is between word-initial voiceless unaspirated stops and prevoiced stops (i.e., voicing begins before the stop closure is released) (Cohen *et al.*, 1961; Keating, 1984; Rietveld and Van Heuven, 2009, pp. 134–135). van Alphen and Smits (2004) investigated the voicing distinction in Dutch initial bilabial and alveolar plosives and showed that speakers indeed generally implemented the contrast in terms of the presence or absence of prevoicing (negative VOT). Perceptually, prevoicing was also the strongest cue that listeners used to classify plosives as voiced or voiceless. In addition to prevoicing, a range of secondary cues (e.g., the burst duration, the intensity of the burst) appeared to play some role in the contrast. Interestingly, some evidence provided by van Alphen and Smits (2004) suggests that voiced stops might not always show full voicing during the closure. We will return to this issue in the discussion section. This study aims to contribute evidence from a true voicing language, by describing CF0 effects after word-initial bilabial stops (/b/ and /p/) in Dutch.

In addition to studying Dutch stops, we aimed to give for the first time a full account of CF0 effects after Dutch fricatives. Standard Dutch is traditionally described as having a phonological distinction between voiced and voiceless fricatives (Booij, 1995). The primary cue for the voiced/voiceless distinction is the presence or the absence of vocal cord vibration in the fricative (Slis and Cohen, 1969; van den Berg, 1988). Moreover, voiceless fricatives tend to be longer and louder than their voiced counterparts (Slis and van Heugten, 1989; Kissine *et al.*, 2003). However, research in the last few decades has provided ample evidence that initial voiced fricatives in standard Dutch are increasingly produced as voiceless because of an ongoing sound change (see next section for more details). The current study will therefore include initial labiodental fricatives /v/ and /f/.

E. Sound change in Dutch fricatives and the question of tonogenesis

The study of CF0 effects is not only interesting from the phonetic and phonological points of view, but it can also tell us something about models of sound change (Hanson, 2009). Lexical tone in many unrelated languages, for instance, is believed to have derived from voicing contrasts in obstruents. Indeed, one of the potential origins for this process (often called tonogenesis) is when vowel F0 differences become the primary cue (Hombert *et al.*, 1979; Kingston, 2011).

Tonogenesis induced by F0 effects is characteristic of the Sino-Tibetan, Hmong, Tai, and Kam-Sui families of languages in East and Southeast Asia (Kingston, 2011). A tone system induced by an earlier [voice] contrast in initial consonants has been illustrated for Vietnamese (Haudricourt, 1954), Seoul Korean (Silva, 2006), but many other examples can be found in this linguistic area (Matisoff, 1973; Mazaudon, 1977). A similar ongoing process has been reported for Afrikaans (Coetzee *et al.*, 2018). Traditionally, prevoiced and voiceless unaspirated plosives ([b]–[p], [t]–[d]) in Afrikaans are produced with comparable VOTs. Yet, the voice contrast is not neutralized but is rather maintained through the F0 contour of the post-stop vowels. Coetzee *et al.* (2018) showed that F0 appeared 50 to 100 Hz lower after phonologically voiced than after phonologically voiceless plosives. The F0 difference continues through at least 70% of the vowel. They therefore concluded that this difference is getting phonologized, resulting in an emergent two-way tonal system.

As mentioned in Sec. I D, a comparable sound change involving the loss of the voicing contrast has extensively been reported for Dutch initial fricatives. There is extensive evidence that Dutch initial voiced fricatives are often produced as voiceless because of a change in voicing, duration, and intensity (Goossens, 1974; Gussenhoven, 1999; Gussenhoven and Bremmer, 1983; Kissine *et al.*, 2003, 2005; Pinget, 2015; Pinget *et al.*, 2020; Van de Velde *et al.*, 1996). There is thus a large consensus that Dutch fricative devoicing is a sound change in progress, but not all fricatives are equally advanced in the process. Labiodental fricatives /f, v/ which are taken into consideration in this study take an in-between position in the change: their devoicing is less advanced than in dorsal fricatives, but further than in alveolar fricatives. Moreover, regional differences were often observed within the Dutch language area as to how far along the sound change process is. Kissine *et al.* (2003, 2005), Van de Velde (1996), and Pinget *et al.* (2020) showed that some regions such as the center and north of the Netherlands are more advanced in the sound change than southern regions.

The devoicing process in Dutch fricatives offers an interesting case to study CF0 effects. First, it allows us to investigate the covariation patterns between voicing cues within the fricatives and post-fricative vowel F0. Specifically, we will be able to compare F0 patterns following highly devoiced /v/ (produced by speakers taking part in the sound change) and following fully voiced /v/ [produced by speakers who are not (yet) taking part in the sound change]. Second, the examination of F0 contours after devoiced fricatives allows us to inspect whether Dutch, a language related to Afrikaans (Reenen and Coetzee, 1996), also shows emerging tonogenesis. Evidence for the maintenance of low onset F0 when [+voice] fricatives are produced phonetically devoiced would point towards this direction. If F0 differences disappear along with voicing, however, then the sound change could result in a merger (Labov, 1994).

F. Summarizing goals and hypotheses

H1: In general, we expect higher F0 values at the onset of a vowel following a voiceless obstruent (/p, f/) than following a voiced obstruent (/b, v/). When compared to a nasal /m/ baseline, F0 after /p, f/ is expected to start higher (i.e., to be raised) at the vowel onset, while F0 after voiced obstruents (/b, v/) is expected to be comparable to the nasal baseline. Such evidence would be taken in favor of an articulatory account of CF0 effects, where voiceless obstruents tend to raise F0 because of stiffening of the vocal folds involved in the production of vocal fold vibration (following, a.o., Halle and Stevens, 1971; Hanson, 2009).

H2: For Dutch voiced fricatives /v/—as they have entered a process of devoicing—we distinguished between three possible scenarios of covariation patterns, i.e., maintenance (H2a), enhancement (H2b), or loss of CF0 effects (H2c), related to the amount of voicing.

H2a: CF0 effects could be maintained: F0 at onset is comparable after a highly devoiced /v/ to after a (normally) voiced /v/. In this case, we thus expect no interaction between the degree of voicing in /v/ and F0 at vowel onset. Such a finding would not provide evidence against an articulatory account of CF0 effects. Indeed, the suppression (or at least diminution) of laryngeal articulatory gestures does not necessarily lead to any consequence in F0 levels. However, such evidence is possibly more in favor of a perceptual enhancement account. While consonantal voicing cues are getting lost (voicing, duration, intensity), F0 is maintained to ensure a perceptual contrast between voiced and voiceless consonants.

H2b: CF0 effects are enhanced in devoiced /v/. In this case, we thus expect an interaction between the degree of voicing in /v/ and F0 at vowel onset: F0 at onset is lower after a highly devoiced /v/ than after a (normally) voiced /v/. Such a correlation between fricative voicing and F0 would form clear evidence for a perceptual enhancement account. Pushing the reasoning of Kingston and Diehl (1994) one step further, it is possible that—because F0 is used as an enhancing gesture—the magnitude of F0 effects is increased by the speaker when fricative voicing is getting weak. CF0 effects in devoiced cases are therefore not only maintained but amplified, and presumably in proportion to the amount of devoicing. This is a possible route towards tonogenesis in Dutch.

H2c: CF0 effects are getting lost in devoiced /v/. Here, we thus expect an interaction between the degree of voicing in /v/ and F0 at vowel onset, but—in this case—F0 at onset is higher (more raised) after a highly devoiced /v/ (and therefore compares more to F0 after /f/) than after a normally voiced /v/. Such a correlation between fricative voicing and F0 would clearly be in favor of the articulatory account: the more the laryngeal gestures disappear, the more F0 effects are getting lost alongside. Specifically, we follow the argument by Löfqvist *et al.* (1989) that the higher crycothyroid muscle activity for voiceless obstruents increases the tension along the vocal folds, and that this higher longitudinal tension inhibits (or helps to inhibit) voicing/phonation. Once phonation does start for the vowel following the consonant, however, the higher longitudinal tension then results in a higher F0 (relative to the same vowel following a voiced consonant). Thus, if a /v/ is produced as more voiceless, presumably involving more crycothyroid muscle activity and presumably with more longitudinal tension along the vocal folds in order to inhibit phonation during the consonant, then we expect a more raised F0 in the beginning of the following vowel.

In the next section, we describe the method for the study. Section III presents the results. Finally, the results are discussed, and the findings are summarized in Sec. IV.

II. METHODS

A. Speech materials and recordings

The data reported here are based on the laboratory speech recordings of test sentences read aloud by 100 native speakers of standard Dutch. We first created 45 Dutch consonant-vowel-consonant (CVC) nonwords, formed by combining an initial target consonant (/v/, /f/, /b/, /p/, or /m/, the vowel /i/ and a final consonant (/g, f, k, l, m, n, p, s, t/) (see the Appendix, Table I). Vowel /i/ was selected because it shows the least regional variation in Standard Dutch (Van der Harst, 2011:159) (i.e., other vowels might have shown significant regional differences in formants) and /i/ was kept constant in the study as we know that the magnitude of the F0 perturbations may vary with the vowel (Ladd and Silverman, 1984; Silverman, 1986; Whalen *et al.*, 1999). The target words were inserted in the focus position of the carrier sentence "Ik neem de ___" [I take the ___]. The use of a carrier sentence and a controlled experiment presentation was meant to force participants to produce every target word with similar intonation patterns, as Hanson (2009) has shown that F0 effects significantly depend on the pitch environment. Participants received no specific instruction regarding the target intonation pattern at the sentence level, but—because the carrier sentences were declarative with a different direct object every trial—they tended to produce the sentence with a narrow focus on the direct object.

A total of 72 sentences (containing 45 targets and 27 fillers) was randomly presented one at a time to each participant on a laptop operating with Linux using Zep (Veenker, 2012). Participants were instructed to read the word aloud in a normal speaking voice as it appeared on the screen. Digital recordings were made at 48 kHz, 24 bits, using an AKG C420 cardioid condenser head-mounted microphone (AKG, Vienna, Austria). This equipment was designed for portability (recording were made in five different labs), while still providing excellent recordings.

B. Speakers

A hundred native speakers of standard Dutch were recruited for this study. They were born and raised in five regions of the Dutch language area, viz. Groningen, South-Holland, Netherlands Limburg, Flemish-Brabant, and West-

Flanders. This selection of regions enables us to cover different stages of devoicing in voiced fricatives, as we know that the change is currently spreading regionally (Kissine *et al.*, 2003; 2005, Pinget, 2015; Pinget *et al.*, 2020). It consists of two regions with speakers typically producing highly devoiced fricatives, such as Groningen and South-Holland, two regions with only incipient stages of devoicing (such as West-Flanders and Flemish-Brabant), and a region with an in-between status in the process of devoicing (see Pinget *et al.*, 2020 for more details).

Of each region, ten males and ten females were recorded. All speakers were highly educated young adults, aged between 18 and 28 years of age (mean = 22.03 years) and attending or recently graduated from a university or a university of applied science. All participants reported having normal hearing and no speech or language disorders. They were paid for their participation.

C. Segmentation and measurements

Realizations of the target obstruents and subsequent vowels were segmented and labeled using PRAAT version 6.1.38 (Boersma and Weenink, 2021). *Fricatives* were segmented by assessing their center of gravity (Gordon *et al.*, 2002), following a segmentation protocol for Dutch (van Son, 2001) (see also Pinget *et al.*, 2020, for details). Following Kissine *et al.* (2003), F0 in the fricative segment was measured with intervals of 10 ms constrained to the 50–400 Hz range. The number of measurements with the presence of F0 was divided by the total number of measurements, multiplied by 100. The resulting relative voicing score ranges from 0 (no voicing throughout the obstruent) to 100 (voicing throughout the obstruent).

Stops were segmented by defining their onset and their offset. The criterion for indicating the onset of the stop is the point where the second formant ($F2$) disappears, which is equivalent to the offset of the preceding vowel (Cho and McQueen, 2005). The criterion for indicating the offset of the stop is the $F2$ onset of the next segment, following Rojczyk (2011, p. 117). The stop onset, offset, and release burst were determined by visual analysis of the spectrographic display and waveforms. Subsequently, VOT was measured as the duration between the stop release and the onset of periodic voicing. As in the study by van Alphen and Smits (2004), our voiced stops /b/ did not always show full voicing during the closure phase. In total, 223 of the 900 realizations of /b/ (i.e., 24.78%) were produced with some (short) cessation of phonation. These devoiced instances were equally often encountered in all regions, but not in all speakers. We will return to this issue in the discussion section.

For the *nasal* onsets and offsets, we looked for changes in the waveform and spectrogram associated with the release of the supraglottal constriction. For offsets specifically, changes were indicated by formants higher than $F1$ being more strongly excited, i.e., higher frequency modulations on the $F1$ oscillations. In the same way, vowel onsets and offsets were aligned with changes in the supraglottal constriction as identified by observations on the waveform and spectrogram. Specifically, vowel offset was defined as the last pitch cycle before a drop in the amplitude or before friction or nasality onset depending on word offset consonant.

Vowel F0 measurements were taken using PRAAT at 11 equidistant time points within the vowels. F0 was constrained to the 100–500 Hz range for females, and to the 75–300 Hz range for males. F0 measures were examined visually and checked by hand. Because we are studying the time course of F0, and because the hypothesized CF0 effects unfold in real time rather than relative time, we chose to transform the relative time points to absolute time in ms after vowel onset (using the relative time points and the duration of each vowel token).

For each individual speaker, the F0 values measured in the first 20 ms of vowels following /m/ were averaged, yielding a speaker-specific vowel-onset F0 centroid in Hz. All F0 measures of a speaker were subsequently expressed in *semitones* relative to this speaker-specific F0 centroid (cf. Shultz *et al.*, 2012). This centering increases the sensitivity of the statistical models to detect possible transitory effects of onset voicing on F0, and the semitone scale reduces the correlation between speaker mean and speaker standard deviation in F0.

Regarding F0 as a function of absolute (not relative) time may introduce the risk, however, that the rare observations at very late absolute time points (from the rarely occurring long vowel tokens) might exert an overdue influence on our statistical models of F0 at vowel onsets. Therefore, we censored the time variable by discarding all observations with absolute time points beyond 194 ms. [The cutoff value of 194 ms was chosen because it was the 99% percentile of the durations of all vowel tokens. This censoring removed 70 (0.1%) of the F0 measurements, which originated from 45 (1%) of the vowel tokens. In addition, 88 observations coming from eight vowel tokens were also discarded because the duration of the vowel token was missing.] Moreover, the observations at time points between 97 and 194 ms were weighted gradually lighter, from 1 to 0 by linear interpolation, again to avoid overdue influence of these observations. (The cutoff value of 97 ms was chosen here too because it was the median of the durations of all vowel tokens; this weighting affected 5810 (11.8%) of the remaining observations, from 2236 (49.7%) of the remaining vowel tokens.) The remaining data set contained 49 342 observations from 4492 vowel tokens.

D. Statistical analyses

The resulting F0 data in semitones were analyzed by means of GAMM (Wood, 2017, Wieling, 2018) using the *mcgv* and *itsadug* packages (Wood, 2017; van Rij *et al.*, 2022) in R (R Core Team, 2022). Three analyses were performed (see Data Availability). In the first GAMM, the speaker-centered F0 values of the remaining weighted observations were regressed onto time (in ms, counted from vowel onset), with different smoothing functions for each initial consonant and for each speaker [f0centered ∼ s (voweltime, by = consonant, k = 20) + consonant + ti (voweltime, talker, bs = "fs," m = 1), see supplementary materials[2]]. To stabilize the GAMM, 20 suspect F0 observations (having >6 semitone difference between adjacent measurements in the same vowel token) from seven vowel tokens (by seven different speakers) were ignored. This resulting GAMM and the subsequent ones reported below also contained a term for autocorrelation of the residuals.

In the second and third analysis, we also included the degree of voicing of the onset consonant as a predictor, but this was done differently for fricative and plosive consonants, yielding two additional GAMMs. For *fricative* consonants /v/ and /f/, the relative degree of voicing (see Sec. II C) and its interaction with smoothed time functions were included in the second GAMM. For *plosive* consonants /b/ and /p/, VOT (in ms) and its interaction with smoothed time functions were included in the third GAMM.

III. RESULTS

A. General model

As described above, a generalized additive mixed model (GAMM) was fit to capture the smoothed effects of time after vowel onset (for each onset consonant separately) on the F0 contour in the critical vowel, allowing for random variation between individual speakers in these smooth functions. Diagnostics of the GAMM did not suggest any problems. A comparison by likelihoods (Bartoń, 2023) of the general GAMM and a null model without any random effects of talkers indicated that in the general GAMM, 48% of the variance in the centered F0 values was due to random effects of talkers (adjusted pseudo R^2 0.50). Figure 1 visualizes the non-random part of this general GAMM as fitted F0 contours, separately for each onset consonant; our interpretations below are driven mainly by this graphical summary of the model. (The parametric and non-parametric estimates of this general GAMM are provided in the Appendix, Table II.)

Figure 1 shows several interesting patterns. First, compared to the baseline condition of F0 after /m/, the F0 after

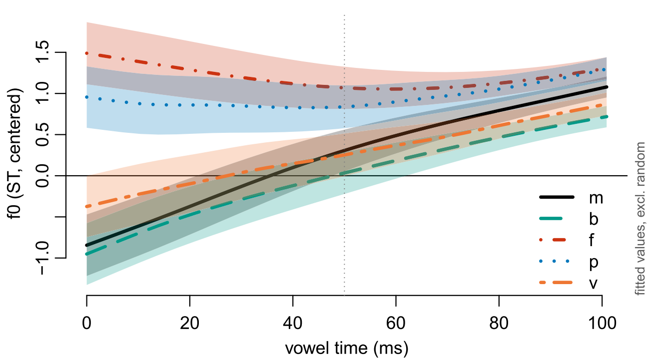


FIG. 1. (Color online) Fitted contours of F0 (in semitones, centered by speaker) over time (in ms from vowel onset) as fitted by nonlinear smooth functions, broken down by onset consonant, with 95% confidence intervals. Random effects are ignored in the fitted F0 values.

unvoiced consonants /f, p/ starts relatively higher at vowel onset (by about 2.0 and 1.5 semitones, respectively, see Fig. 1). In contrast, second, the F0 after **voiced** consonants /v, b/ does not differ from that after /m/. At vowel onset, the confidence bands of F0 after unvoiced consonants /b, p/ are at least one semitone higher than those after voiced /v, f/. Third, these voicing effects of the onset consonant on F0 are only transitory, and are no longer noticeable after approximately 50 ms in the vowel (when the 95% confidence intervals of /m/ and /p/ start to overlap, see Fig. 1).

Because consonant voicing manifests itself in different ways for plosives (VOT) than it does for fricatives (degree of voicing during frication), we will further inspect the F0 contours in separate GAMMs for onset plosives and onset fricatives, respectively (see Sec. II D). The ongoing change in voicing of Dutch initial fricatives—but not in initial plosives—also suggests that different sociophonetic processes may be at play in the two classes of tokens.

B. Additional model for F0 after fricatives

As described above, a second GAMM was fit to capture the smoothed effects of time after vowel onset (for each onset consonant separately) on the F0 contour in the critical vowel, as well as the smoothed effects of the relative amount of voicing of the fricative, as well as their interaction, while allowing for random variation between individual speakers in the main effects of time. Diagnostics of this GAMM did not suggest any problems. According to a Likelihood Ratio Test, this complex model including the interactions of time and relative degree of voicing performed considerably better than a simpler model with main effects, random intercepts, and slopes, but without interaction terms [$\chi^2(73) = 384$, $p < 0.0001$]. In Fig. 2, the gradient effect of the degree of consonant voicing on the F0 contour of the following vowel is visualized as a contour plot, showing the joint effects of time (along X, first part of vowel

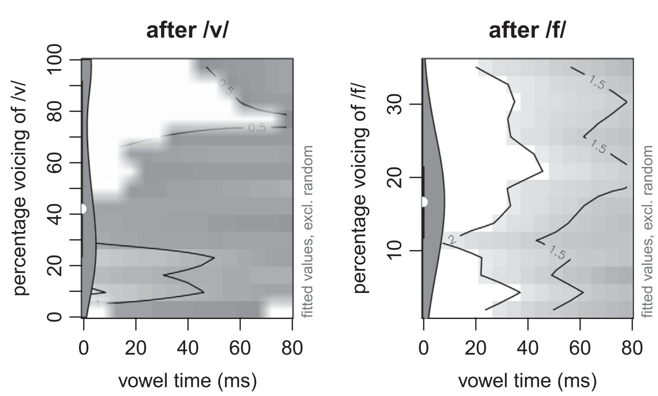


FIG. 2. Contour plots showing the interaction between vowel time (in ms, first part of vowel only) and the degree of voicing of the onset consonant (in percentage points, rescaled), on F0 (in semitones, centered, lighter shade indicates higher F0). The lefthand panel shows the interaction for vowel tokens after /v/; the righthand panel shows the interaction after /f/; voicing axes are scaled differently across panels for clarity. The density of the degree of voicing is shown along the left axis (with quartiles and median). Random effects have been ignored in the fitted F0 values. White areas have essentially no interaction.

JASA
https://doi.org/10.1121/10.0021070

only) and degree of voicing (scaled, along *Y*) on F0 (*Z*, higher F0 in lighter shade of gray).

Figure 2 shows a clear gradient effect of the relative degree of voicing, in particular at the beginning of the vowel after /v/. If an onset/v/ is less voiced, then F0 at vowel onset tends to start higher (lefthand panel, lower left, lighter shade) and decrease more. These gradient effects of voicing of /v/ are only present in the first part of the target vowel and disappear after about 40 ms in the vowel. Most of the onset /f/ consonants have a low degree of voicing (righthand panel, about 90% of /f/ tokens are below <25% voiced) and these show the F0 pattern also visible in Fig. 1: a relatively high F0 at vowel onset, followed by a decrease in F0 in the first 60 ms of the vowel. (We ignore the "visible ridge" at 10% voicing of /f/ as a probable overfitting).

C. Additional model for F0 after plosives

As described above, a third GAMM was fit to capture the smoothed effects of time after vowel onset (for each onset plosive consonant separately) on the F0 contour in the critical vowel, as well as the smoothed effects of the VOT of the onset plosive, as well as their interaction, while allowing for random variation between individual speakers in the main effects of time. Diagnostics of this GAMM did not suggest any problems. According to a Likelihood Ratio Test, this complex model including the interactions of time and relative degree of voicing did not perform significantly better than a simpler model with main effects, random intercepts, and slopes, but without interaction terms [$\chi^2(48) = 63.4$, $p = 0.063$]. This suggests that the interaction did not significantly contribute to the complex GAMM for vowels following plosives. Figure 3, visualizing the gradient effect of the degree of consonant VOT on the F0 contour of the following vowel, confirms this absence of interaction. No interactions are visible in the onset of the vowel in the areas with higher density along the VOT axis.

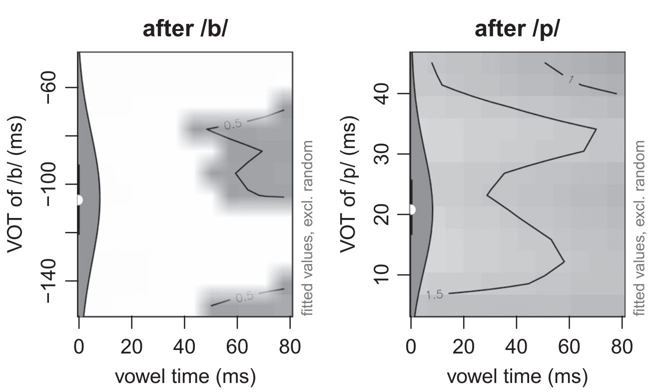


FIG. 3. Contour plots showing the interaction between vowel time (in ms, first part of vowel only) and VOT (in ms), on F0 (in semitones, centered, lighter shade indicates higher F0). The lefthand panel shows the interaction for vowel tokens after /b/; the righthand panel shows the interaction after /p/; VOT axes are scaled differently across panels for clarity. The density of VOTs is shown along the left axis (with quartiles and median). Random effects have been ignored in the fitted F0 values. White areas have essentially no interaction.

IV. GENERAL DISCUSSION

In this study, we examined the CF0 after Dutch obstruents. Our results showed that F0 was raised after voiceless obstruents, while it remained largely unaffected after voiced obstruents, where the F0 contour did not differ from that after the nasal consonant /m/ regarded as baseline. The results are discussed in relation to the hypothesized source of CF0 perturbations (Sec. IV A). The covariation patterns and the differences between fricatives and stops are discussed in Secs. IV A and IV B, while Sec. IV C reviews the time course, the magnitude, and the contours of the effects.

A. F0 perturbations: Articulatory vs auditory account

First, we asked whether CF0 perturbations are best accounted for by an automatic, articulatory source or an auditory enhanced source. For fricative /v/, it was shown that fricatives devoicing and F0 level tend to covary: the less voicing in /v/, the higher F0 at onset. In other words, F0 at onset is higher (more raised) after a highly devoiced /v/ (and therefore compares more to F0 after /f/) than after a (normally) voiced /v/. It seems thus that CF0 difference is decreasing because of the ongoing sound change of /v/. This result is clearly in favor of the articulatory account: the more the laryngeal gestures disappear, the more F0 effects are getting lost alongside. The lowered F0 following voiced fricatives does not become phonologized, showing that the ongoing sound change in Dutch will most probably result in a full /v, f/ merger, and not in tonogenesis.

For bilabial stops, however, our results did not reveal any covariation patterns between VOT and F0 levels. Kirby and Ladd (2015) did find an inverse correlation between voicing lead and onset F0: i.e., the longer prevoicing was accompanied by a lower F0 at the onset of the following vowel. This finding was, however, not replicated in Kirby and Ladd (2016), a follow up study on the same languages where obstruents in carrier phrase contexts were analysed instead of words in isolation. The examination of covariation patterns in stops in the present study does not provide enough evidence in favor of either an automatic, articulatory source or an auditory enhanced source of CF0 effects. In the next section, possible reasons for this lack of co-variation patterns in stops are discussed.

Second, we asked whether CF0 perturbations in the case of an articulatory mechanism were better described in terms of a lowered F0 after voiced obstruents or a raised F0 after voiceless obstruents (or possibly both mechanisms at the same time). The covariation patterns do not directly provide evidence on this matter, but the nasal baseline allows us to disentangle whether F0 was raised or F0 lowered. The results clearly showed that F0 was higher after voiceless obstruents than after voiced obstruents and after /m/. F0 following voiced stops was not lowered relative to that of /m/. Therefore, our data lend more support to the effect of F0 raising of voiceless obstruents as the primary source of CF0 perturbations. This finding corresponds to the evidence provided by Gao and Arai (2019) for Japanese, and by Hanson

(2009) for English; however, it contradicts the results for Afrikaans by Coetzee *et al.* (2018), who found lowering of F0 after voiced, next to effect of F0 raising of voiceless obstruents. We could not observe any intended gesture to enhance the saliency of the [+voice] feature as compared to the nasal baseline, as was proposed by Kingston and Diehl (1994). With Hanson (2009), we interpret these findings as being the result of an increase in active vocal fold stiffening during the voiceless obstruent consonants that carries over to the following vowel as originally proposed by Halle and Stevens (1971) and Löfqvist *et al.* (1989). In short, these data provide clear support for articulatory effects that fall out from gestural overlap between a voiceless obstruent and the following vowel.

B. No covariation pattern in stops

This study fails to reveal any covariation patterns between VOT and F0 levels in Dutch stops. This has been the case in different previous studies on different languages. Dmitrieva *et al.* (2015), for instance, have reported within-category VOT and onset F0 to be uncorrelated in Spanish and English data. Kirby and Ladd (2016) could not replicate an inverse correlation between voicing lead and onset F0 (i.e., the longer prevoicing accompanied by a lower F0 at the onset of the following vowel) they found in 2015, once they completed a follow up study on French and Italian obstruents. They also reported considerable individual variation in the strength and direction of correlations, a finding also reported by Dmitrieva *et al.* (2015) and Gao and Arai (2019). The current study presented a very large pool of individual speakers ($N = 100$, balanced for gender, and born and raised in different regions), and at the same time accounted for individual variation in modeling F0 contours by means of GAMM, thus yielding better generalizable F0 contours (Baayen, 2008, Sec. 7.3). It thus seems that—once individual effects are properly modelled in the analyses—no covariation patterns can be found in the stop data.

A possible explanation for our failure to find covariation patterns in stops in the current study (and the previous studies) could be the lack of variation within category. Despite the substantial amount of data in the current study (a hundred speakers, nine tokens per onset consonant), the homogeneity in stop production was very large. In contrast, there was a lot of variation within the /v/ category. As mentioned in the introduction, Dutch voiced fricatives including labiodentals have been involved in the past decennia in a devoicing process for which we have strong evidence. The large collection of data including a hundred speakers from different regions of the Dutch-speaking area and with different degrees of devoicing succeeded in eliciting very heterogeneous realizations of Dutch /v/, which possibly explains why a significant covariation pattern could be found for this category. The lack of variation within the category could be a reason why there was not a significant effect in the other three categories (/f, b, p/).

Another interesting thought in the case of Dutch stops is that there was in this study a number of voiced stops showing a short cessation of phonation that leads to a short period of voicelessness in the prevoicing. Other studies considering Dutch have reported such a phenomenon where vocal pulsing dropped before the release or cases where vocal pulsing ceases in the middle of the prevoiced part and started again before the burst (Foulkes *et al.*, 2010; van Alphen and Smits 2004; Pinget, 2015; Pinget *et al.*, 2020). van Alphen and Smits (2004), for instance, found that only 75% of word-initial "voiced" stops produced when reading a wordlist had voicing during closure. Pinget (2015) further investigated this apparent cessation of phonation during closure and confirmed that Dutch /b/ might not always be produced as fully voiced. Pinget *et al.* (2020) examined this "devoicing" in bilabial stops as a possible case of incipient sound change and showed that only some individuals in a large pool of participants clearly devoiced initial voiced stops, but it does not yet form a strong community pattern. Recent findings by Pfiffner (2021) also showed individual speakers of Dutch with devoicing. These speakers with the largest devoicing rates were all in the younger age groups (22–29 years old). This suggests that Dutch [+voice] stops might not always show continuous voicing during the closure, and consequently analyzing VOT as the traditional primary cue in stop might not fully cover what is happening there. In the current data set, 25% of the voiced stops were to some extent devoiced, a proportion comparable to van Alphen and Smits (2004). Taken together, evidence in these studies on Dutch stops could potentially point towards stop devoicing, but it is currently unclear whether plosive devoicing is a stable variation or the beginning of a sound change. This is of interest because some scholars have proposed that it is the phonological, rather than the phonetic status of stops that triggers F0 raising or lowering in the following vowel (Ohde, 1984; Hanson, 2009; Dmitrieva *et al.*, 2015). The study of English and Spanish (a true voicing language) stops by Dmitrieva *et al.* (2015), for instance, showed that F0 contours in both languages were highly similar, and therefore predictable from the phonological status of the stops, and not from the specific phonetic realization of VOT. Dutch stops showing some cessation of phonation might provide additional evidence supporting this claim, as there is variation within the [+voice] category. This pattern merits to be followed up as it can help to provide additional evidence in the future examination of Dutch stops and to possibly reconsider the "aspirating" vs "voicing" dichotomy either as a continuum or as a multidimensional space where more cues are involved than just VOT.

C. Time course, magnitude, and contours of CF0 perturbations

The GAMM models presented in the current study allowed us to examine F0 perturbations over the vowel duration in order to get insight into the time course of F0 effects. The results showed that the F0 effect not only involved F0 values immediately after the consonants, but it entailed an

F0 movement that extended further into the vowel, mostly into its first half. In that respect, Dutch seems to be different than Cantonese measured by Francis *et al*. (2006) where the reported F0 effects were shown to be limited to the very onset of the vowels (0–10 ms). Our findings correspond more to results by Hanson (2009) and Kenstowicz and Park (2006), reporting perturbations at least until midvowel.

In languages (possibly) undergoing tonogenesis, the magnitude of the F0 difference was often shown to be much greater than in languages where there is no tonogenesis: 45–60 Hz (3.3 to 4.5 semitones) in Seoul Korean (Silva, 2006), about 40 Hz (over three semitones) in young female speakers of Afrikaans (Coetzee *et al*., 2018) vs 8–16 Hz for English, German, and French (around one semitone) (reported by Coetzee *et al*., 2018). Gao and Arai (2019) also showed that onset F0 effects are enhanced in Tokyo Japanese in contexts where the VOT cue was not sufficiently reliable for potential tonal development. In the current study, the magnitude of F0 perturbations turned out to be at least one semitone between voiced and voiceless obstruents at vowel onset, which is well above the just noticeable difference for frequency differences (of about 1%, or 0.2 semitone; Fastl and Zwicker, 2007, Sec. 7.2; Lehiste, 1970, p. 64). As far as Dutch is concerned, Slis and Cohen (1969) measured F0 values in naturally produced /Ca/ syllables and reported that the "top frequency" in the vowel following voiceless consonants was on average 6 Hz higher (i.e., presumably about 1 semitone, their Fig. 13, p. 99) than following voiced consonants. van Alphen and Smits (2004) measured a difference of 16 Hz (about 1.6 semitone) between voiced and voiceless labial and alveolar stops without providing any distinction between male and female speakers. Pfiffner (2021) recently reported larger F0 differences (19–56 Hz, or 2.5–4.7 semitones). The latter three studies, however, only included words recorded in isolation. The magnitude of F0 differences cannot directly be compared across studies as we know that it highly depends on the general intonational context, the specific vowels, etc. (Hanson, 2009; Kirby and Ladd, 2016; Gao and Arai, 2019). It is possible that the high front vowel /i:/ used in this study has led to relatively large F0 effects. Hence, a replication of comparable F0 effects with other Dutch vowels is desirable in future work. Importantly, the current recordings of obstruents were made within a high pitch context (i.e., words in focus position within a carrier sentence) which does reflect natural speech situations better than words in isolation, and we have expressed F0 in semitones to allow for better future comparisons of the magnitude of the effects.

Note that there were interesting differences in the (magnitude of) CF0 effect between stops and fricatives. Overall, CF0 effects seemed larger for fricatives than for stops: F0 was higher after /f/ compared to the baseline /m/ than after /p/. Moreover, the CF0 effects after /f/ were observed to last longer into the vowel than after /p/. On the one hand, such a finding might not be expected from an automatic, articulatory point of view. Indeed, we would expect the effects to be smaller in fricatives because the articulatory constraints on voicing are not as great in fricatives compared to plosives. On the other hand, it has often been reported that voiceless fricatives have a larger glottal opening than stops (e.g., Collier *et al*., 1979), which could lead to a higher f0 at the onset of the vowel and possibly explain the magnitude difference. Because of the great overlap in confidence bands, it is difficult to draw clear conclusions, but it might be worth to address this issue in future studies.

The comparison between F0 differences for Dutch in the current study (and their magnitude) and the Afrikaans data provided by Coetzee *et al*. (2018) allows us to raise the question of what drives this type of variation in the first place. Since Dutch and Afrikaans are closely related, also in terms of phoneme inventories and distribution (see De Villiers and Ponelis, 1992; Ponelis, 1993, for more details), it seems improbable that differences in functional load have triggered the development of a tone contrast in Afrikaans (in order to compensate for the loss of voicing,) and not in Dutch. A further question then arises: What is the role of language contact here, since Afrikaans is surrounded by Bantu tone languages? A future, thorough examination of F0 effects in Dutch and Afrikaans with a consistent methodology has the potential to offer crucial insight in the process of tonogenesis and the role of language contact.

As far as the F0 contours are concerned, the GAMM showed a rising F0 pattern after /b/, similar to the pattern after the baseline level /m/ (see Fig. 1), which corresponds to findings by Hanson (2009). As summarized in the previous section, the F0 contours after /v/ were highly dependent on the amount of voicing. We observed no clear evidence of falling F0 patterns after voiceless obstruents. We suggest along with Hanson (2009) that F0 contour shapes could be considered less relevant than other CF0 aspects (such as magnitude and time course) as the displayed patterns are highly dependent on the pitch environments. Nevertheless, it would be insightful to investigate which contours and slopes are perceptually salient in which contexts, along the lines proposed by Serniclaes (1986) and Silverman (1986). While the duration and the magnitude of CF0 perturbations are large enough to be perceptually salient, such CF0 differences are typically neutralized or compensated for during perception ('t Hart *et al*., 1990; Peterson, 1986; Silverman, 1986). The question then remains why the perceptual system sometimes fails to compensate for these CF0 differences, so that these CF0 differences may eventually become phonologized. Well-controlled perceptual experiments are necessary to gain a better understanding of the perceptual relevance of F0 differences and their cue weighting patterns.

V. CONCLUSION

This study examined the consonant-related F0 perturbations in Dutch after initial fricatives (/v, f/) and stops (/b, p/), as compared to after the nasal /m/. Our results showed that F0 was raised after voiceless obstruents, while it remained largely unaffected after voiced obstruents, where

https://doi.org/10.1121/10.0021070

the F0 contour did not differ from that after the nasal consonant /m/ regarded as baseline.

F0 patterns after Dutch initial /v/ were particularly interesting because large individual variation across speakers was observed in the realization of voicing: some fricatives are highly devoiced (as the result of an ongoing sound change) and others kept fully voiced. Fricative voicing in /v/ and F0 level tend to covary: the less voicing in /v/, the higher F0 at onset. There was no trace of an active gesture to explicitly lower pitch after highly devoiced fricatives, as would be predicted by an auditory account. In conclusion, F0 perturbations after Dutch obstruents and their covariation patterns are taken as additional evidence to support an articulatory cause of consonant-related F0 effects. Furthermore, Dutch was shown to be an interesting case study for further studies on the production and perceptual nature of consonant-related F0 effects and their relation to voicing contrasts.

ACKNOWLEDGMENTS

We are very grateful to the anonymous reviewers for helping to clarify and improve this work. We thank the following institutes for providing recording facilities: the Department of Dutch Linguistics of Ghent University, the Leiden University Phonetics Laboratory, 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 Stefano Coretta, Joseph Casillas, Timo Roettger, Martijn Wieling, and Cesko Voeten for their expert statistical guidance. Our interpretation of the results benefited from discussions with many colleagues, of whom we especially thank James Kirby, Jiayin Gao, René Kager, and Patrice Beddor. The data collection for this study was supported by the Netherlands Organization for Scientific Research (NWO Grant No. 322-75-002).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts of interest to disclose. The data were collected in accordance with institutional and national regulations in effect at the time of data collection (2013), and informed consent was obtained from all participants in accordance with ASA requirements.

TABLE I. List of stimulus words, in Dutch orthography.

/b/	/p/	/v/	/f/	/m/
bieg	pieg	vieg	fieg	mieg
bief	pief	vief	fief	mief
biek	piek	viek	fiek	miek
biel	piel	viel	fiel	miel
biem	piem	viem	fiem	miem
bien	pien	vien	fien	mien
biep	piep	viep	fiep	miep
bies	pies	vies	fies	mies
biet	piet	viet	fiet	miet

TABLE II. Summary of the general GAMM of F0 (Sec. III A), having as predictors the onset consonant (with /m/ as baseline) and vowel time after vowel onset (in ms), with random smooths per speaker. $R^2_{adj} = 0.374$, deviance explained 38.4%, $N = 46\,392$.

Parametric coefficients	Estimate	Standard error	*t* value	Probability
baseline (/m/)	0.945	0.045	20.98	<0.0001
/b/	−0.444	0.054	−8.18	<0.0001
/f/	0.592	0.059	9.99	<0.0001
/p/	2.231	0.056	39.59	<0.0001
/v/	0.801	0.057	14.17	<0.0001
smooth terms	EDF[a]	Ref. EDF	*F* value	Probability
vowtime*/m/	4.4	5.9	32.67	<0.0001
vowtime*/b/	5.5	7.4	25.07	<0.0001
vowtime*/f/	5.5	7.3	7.77	<0.0001
vowtime*/p/	16.4	17.6	136.27	<0.0001
vowtime*/v/	14.2	16.0	54.33	<0.0001

[a]Estimated degrees of freedom (Wood, 2017).

DATA AVAILABILITY

Data, analysis code and supplementary material are available at https://doi.org/10.24416/UU01-2XJTYX.

APPENDIX

A list of the stimulus words (in Dutch orthography) is provided Table I. Table II shows the summary of the general GAMM of F0 (Sec. III A), having as predictors the onset consonant (with /m/ as baseline) and vowel time after vowel onset (in ms), with random smooths per speaker.

[1]For more details on the different Germanic languages: see Wissing (1991) on Afrikaans; Tiersma (1985) on (West) Frisian; Birnbaum (1979) and Katz (1987) on Yiddish; Hutters (1985) on Danish; Thráinsson (1978) on Icelandic; Vanvik (1972) and Kristoffersen (2000) on Norwegian; Jessen (1998) and Jessen and Ringen (2002) on German; Ladd and Schmid (2018) on Swiss German.

[2]See the supplementary material at https://doi.org/10.24416/UU01-2XJTYX.

Baayen, H. R. (**2008**). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R* (Cambridge University Press, Cambridge, UK).

Barton´, K. (**2023**). "MuMIn: Multi-model inference, (R package version 1.47.5)," https://CRAN.R-project.org/package=MuMIn (Last viewed June 11, 2023).

Beckman, J., Jessen, M., and Ringen, C. (**2013**). "Empirical evidence for laryngeal features: Aspirating vs true voice languages," J. Ling. **49**, 259–284.

Birnbaum, S. (**1979**). *Yiddish. A Survey and a Grammar* (Manchester University Press, Manchester, UK).

Boersma, P., and Weenink, D. (**2021**). "PRAAT: Doing phonetics by computer (version 6.1.38) [computer program]," http://www.praat.org/ (Last viewed January 2, 2021).

Booij, G. E. (**1995**). *The Phonology of Dutch* (Clarendon, Oxford, UK).

Chen, Y. (**2011**). "How does phonology guide phonetics in segment–F0 interaction?," J. Phon. **39**, 612–625.

Cho, T., and Ladefoged, P. (**1999**). "Variation and universals in VOT: Evidence from 18 languages," J. Phon. **27**(2), 207–229.

Cho, T., and McQueen, J. M. (**2005**). "Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress," J. Phon. **33**(2), 121–157.

Coetzee, A. W., Beddor, P. S., Shedden, K., Styler, W., and Wissing, D. (**2018**). "Plosive voicing in Afrikaans: Differential cue weighting and tonogenesis," J. Phon. **66**, 185–216.

Cohen, A., Ebeling, C. L., Fokkema, K., and Holk, A. G. F. (**1961**). *Fonologie Van Het Nederlands en Het Fries* (*Phonology of Dutch and Frisian*) (Nijhoff, Leiden, the Netherlands).
Collier, R., Lisker, L., Hirose, H., and Ushijima, T. (**1979**). "Voicing in intervocalic stops and fricatives in Dutch," J. Phon. **7**(4), 357–373.
Davidson, L. (**2016**). "Variability in the implementation of voicing in American English obstruents," J. Phon. **54**, 35–50.
De Villiers, M., and Ponelis, F. A. (**1992**). *Afrikaanse Klankleer (Afrikaans Phonology)* (Tafelberg, Cape Town, South Africa).
Dmitrieva, O., Llanos, F., Shultz, A. A., and Francis, A. L. (**2015**). "Phonological status, not voice onset time, determines the acoustic realization of onset F0 as a secondary voicing cue in Spanish and English," J. Phon. **49**, 77–95.
Fastl, H., and Zwicker, E. (**2007**). *Psychoacoustics: Facts and Models*, 3rd ed. (Springer, Berlin).
Flege, J. E. (**1982**). "Laryngeal timing and phonation onset in utterance-initial English stops," J. Phon. **10**, 177–192.
Foulkes, P., Docherty, G., and Jones, M. (**2010**). "Analysing stops," in *Sociophonetics*, edited by M. Di Paolo and M. Yaeger-Dror (Routledge, Abingdon, UK), pp. 58–71.
Francis, A. L., Ciocca, V., Wong, V. K. M., and Chan, J. K. L. (**2006**). "Is fundamental frequency a cue to aspiration in initial stops?," J. Acoust. Soc. Am. **120**, 2884–2895.
Gao, J., and Arai, T. (**2019**). "Plosive (de-)voicing and F0 perturbations in Tokyo Japanese: Positional variation, cue enhancement, and contrast recovery," J. Phon. **77**, 100932.
Goossens, J. (**1974**). *Historische Phonologie Des NiederläNdisches* (*Historical Phonology of Dutch*) (Niemeyer, Tübingen, Germany).
Gordon, M., Barthmaier, P., and Sands, K. (**2002**). "A cross-linguistic acoustic study of voiceless fricatives," J. Int. Phon. Assoc. **32**(2), 141–174.
Gussenhoven, C. (**1999**). "Illustrations of the IPA: Dutch," in *Handbook of the International Phonetic Association* (Cambridge University Press, Cambridge, UK), pp. 74–77.
Gussenhoven, C., and Bremmer, R. H. (**1983**). "Voiced fricatives in Dutch: Sources and present-day usage," NOWELE **2**, 55–71.
Haggard, M., Stephen, A., and Callow, M. (**1970**). "Pitch as a voicing cue," J. Acoust. Soc. Am. **47**(2B), 613–617.
Haggard, M., Summerfield, Q., and Roberts, M. (**1981**). "Psycho-acoustical and cultural determinants of phoneme boundaries: Evidence from trading F0 cues in the voiced–voiceless distinction," J. Phon. **9**, 49–62.
Halle, M., and Stevens, K. N. (**1971**). "A note on laryngeal features," MIT Quart. Prog. Rep. **101**, 198–212.
Hanson, H. M. (**2009**). "Effects of obstruent consonants on fundamental frequency at vowel onset in English," J. Acoust. Soc. Am. **125**, 425–441.
Haudricourt, A. (**1954**). "De l'origine des tons en Vietnamien" ("On the origin of tone in Vietnamese"), J. Asiatique **242**, 69–82.
Hombert, J. M. (**1978**). "Consonant types, vowel quality, and tone," in *Tone* (Academic Press, New York), pp. 77–111.
Hombert, J.-M., Ohala, J. J., and Ewan, W. G. (**1979**). "Phonetic explanations for the development of tones," Language **55**, 37–58.
Honda, K., Hirai, H., Masaki, S., and Shimada, Y. (**1999**). "Role of vertical larynx movement and cervical lordosis in F0 control," Lang. Speech **42**, 401–411.
House, A. S., and Fairbanks, G. (**1953**). "The influence of consonant environment upon the secondary acoustical characteristics of vowels," J. Acoust. Soc. Am. **25**, 105–113.
Hutters, B. (**1985**). "Vocal fold adjustments in aspirated and unaspirated stops in Danish," Phonetica **42**, 1–24.
Jacewicz, E., Fox, R. A., and Lyle, S. (**2009**). "Variation in stop consonant voicing in two regional varieties of American English," J. Int. Phonetic Assoc. **39**, 313–334.
Jessen, M. (**1998**). *Phonetics and Phonology of Tense and Lax Obstruents in German* (John Benjamins, Amsterdam).
Jessen, M., and Ringen, C. (**2002**). "Laryngeal features in German," Phonology **19**, 189–218.
Katz, D. (**1987**). *Grammar of the Yiddish Language* (Duckworth, London).
Keating, P. A. (**1984**). "Phonetic and phonological representation of stop consonant voicing," Language **60**, 286–319.
Kenstowicz, M., and Park, C. (**2006**). "Laryngeal features and tone in Kyungsang Korean: A phonetic study," Studies Phon. Phonol. Morphol. **12**(2), 247–264.
Kingston, J. (**2007**). "Segmental influences on F0: Automatic or controlled?," in *Tones and Tunes*, edited by C. Gussenhoven and T. Riad (Mouton de Gruyter, Berlin, Germany), Vol. 2, pp. 171–210.
Kingston, J. (**2011**). "Tonogenesis," in *The Blackwell Companion to Phonology*, edited by M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice (Wiley, New York).
Kingston, J., and Diehl, R. L. (**1994**). "Phonetic knowledge," Language **70**, 419–454.
Kirby, J. (**2018**). "Onset pitch perturbations and the cross-linguistic implementation of voicing: Evidence from tonal and non-tonal languages," J. Phon. **71**, 326–354.
Kirby, J., and Ladd, D. R. (**2015**). "Stop voicing and F0 perturbations: Evidence from French and Italian," in *Proceedings of the 18th International Congress of Phonetic Sciences*, August 10–14, Glasgow, UK, Paper No. 740.
Kirby, J. P., and Ladd, D. R. (**2016**). "Effects of obstruent voicing on vowel F0: Evidence from 'true voicing' languages," J. Acoust. Soc. Am. **140**, 2400–2411.
Kissine, M., Van de Velde, H., and van Hout, R. (**2003**). "An acoustic study of standard Dutch /v/, /f/, /z/ and /s/," Linguistics Netherlands **20**(1), 93–104.
Kissine, M., Van de Velde, H., and van Hout, R. (**2005**). "Acoustic contributions to sociolinguistics: Devoicing of /v/ and /z/ in Dutch," Univ. Penn. Work. Papers Linguist. **10**(2), 144–155.
Kristoffersen, G. (**2000**). *The Phonology of Norwegian* (Clarendon, Oxford, UK).
Krug, P. K., Gerazov, B., van Niekerk, D. R., Xu, A., Xu, Y., and Birkholz, P. (**2021**). "Modelling microprosodic effects can lead to an audible improvement in articulatory synthesis," J. Acoust. Soc. Am. **150**(2), 1209–1217.
Labov, W. (**1994**). *Principles of Language Change: Internal Factors* (Blackwell, Oxford, UK), Vol. 1.
Ladd, D. R., and Schmid, S. (**2018**). "Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops," J. Phon. **71**, 229–248.
Ladd, D. R., and Silverman, K. E. A. (**1984**). "Vowel intrinsic pitch in connected speech," Phonetica **41**(1), 31–40.
Lehiste, I. (**1970**). *Suprasegmentals* (MIT Press, Cambridge, MA).
Lehiste, I., and Peterson, G. E. (**1961**). "Some basic considerations on the analysis of intonation," J. Acoust. Soc. Am. **33**, 419–425.
Lisker, L., and Abramson, A. S. (**1964**). "A cross-language study of voicing in initial stops: Acoustical measurements," Word **20**, 384–422.
Löfqvist, A., Baer, T., McGarr, N. S., and Story, R. S. (**1989**). "The cricothyroid muscle in voicing control," J. Acoust. Soc. Am. **85**, 1314–1321.
Löfqvist, A., Koenig, L. L., and McGowan, R. S. (**1995**). "Vocal tract aerodynamics in /aCa/ utterances: Measurements," Speech Commun. **16**(1), 49–66.
Matisoff, J. A. (**1973**). "Tonogenesis in southeast Asia," in *Consonant Types and Tone*, edited by L. M. Hyman (University of Southern California, Los Angeles, CA), pp. 71–95.
Mazaudon, M. (**1977**). "Tibeto-Burman tonogenetics," Linguist. Tibeto-Burman Area **3**, 1–123.
Mohr, B. (**1971**). "Intrinsic variations in speech signal," Phonetica **23**, 65–93.
Ohala, J. J. (**1975**). "Phonetic explanations for nasal sound patterns," in *Nasalfest: Papers from a Symposium on Nasals and Nasalization*, edited by C. A. Ferguson, L. M., Hyman, and J. J. Ohala (Language Universals Project, Stanford, CA), pp. 289–316.
Ohde, R. N. (**1984**). "Fundamental frequency as an acoustic correlate of stop consonant voicing," J. Acoust. Soc. Am. **75**, 224–230.
Petersen, R. N. (**1986**). "Perceptual compensation for segmentally conditioned fundamental frequency perturbation," Phonetica **43**(1–3), 31–42.
Pfiffner, A. M. (**2021**). "Perception and production of [voice] contrasts in Dutch word-initial plosives," Proc. Ling. Soc. Am. **6**(1), 340–353.
Pinget, A. C. H., Kager, R., and Van de Velde, H. (**2020**). "Linking variation in perception and production in sound change: Evidence from Dutch obstruent devoicing," Lang. Speech **63**(3), 660–685.
Pinget, A. F. (**2015**). "The actuation of sound change," Ph.D. thesis, Utrecht University, Utrecht, The Netherlands.
Ponelis, F. A. (**1993**). *The Development of Afrikaans* (Peter Lang, Frankfurt, Germany).
R Core Team (**2022**). "R: A language and environment for statistical computing (version 4.2.2) [computer software]" (R Core Team, Vienna, Austria).

Reenen, P., and Coetzee, A. (**1996**). "Afrikaans, a daughter of Dutch," in *The Origins and Development of Emigrant Languages (Proceedings of the Second Rasmus Colloquium, Odense University, Nov. 1994)*, edited by H. F. Nielsen and L. Schøsler (Odense University Press, Odense, Denmark), pp. 71–101.

Repp, B. H. (**1982**). "Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception," Psychol. Bull. **92**, 81–110.

Rietveld, T., and Van Heuven, V. J. (**2009**). *Algemene Fonetiek* (*General Phonetics*), 3rd ed. (Coutinho, Bussum, the Netherlands).

Rojczyk, A. (**2011**). "Perception of English voice onset time continuum by Polish learners," in *The Acquisition of L2 Phonology*, edited by J. Arabski and A. Wojtaszek (Multilingual Matters, Bristol, UK), pp. 37–58.

Serniclaes, W. (**1986**). "Etude exp érimentale de la perception du trait de voisement des occlusives du français" ("Experimental study of the perception of the voicing feature in french stops"), Ph.D. thesis, Université Libre de Bruxelles, Brussels, Belgium.

Shultz, A. A., Francis, A. L., and Llanos, F. (**2012**). "Differential cue weighting in perception and production of consonant voicing," J. Acoust. Soc. Am. **132**, EL95–EL101.

Silva, D. J. (**2006**). "Acoustic evidence for the emergence of tonal contrast in contemporary Korean," Phonology **23**, 287–308.

Silverman, K. (**1986**). "F_0 segmental cues depend on intonation: The case of the rise after voiced stops," Phonetica **43**(1–3), 76–91.

Slis, I., and Cohen, A. (**1969**). "On the complex regulating the voiced-voiceless distinction I and II," Lang. Speech **12**, 80–102.

Slis, I., and van Heugten, M. (**1989**). "Voiced-voiceless distinction in Dutch fricatives," in *Linguistics in The Netherlands*, edited by H. Bennis and A. van Kemenade (Foris, Dordrecht), pp. 123–132.

't Hart, J., Collier, R., and Cohen, A. (**1990**). *A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody* (Cambridge University Press, Cambridge, UK).

Thráinsson, H. (**1978**). "The phonology of Icelandic preaspiration," Nordic J. Ling. **1**, 3–54.

Tiersma, P. (**1985**). *Frisian Reference Grammar* (Foris, Dordrecht).

Ting, C., Clayards, M., Sonderegger, M., and McAuliffe, M. (**2023**). "The cross-linguistic distribution of vowel and consonant intrinsic F0 effects," PsyArXiv https://doi.org/10.31234/osf.io/64nhs.

van Alphen, P., and Smits, R. (**2004**). "Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing," J. Phon. **32**(4), 455–491.

van den Berg, R. (**1988**). "The perception of voicing in Dutch two-obstruent sequences," Ph.D. thesis, Radboud University, Nijmegen, The Netherlands.

van der Harst, S. (**2011**). "The vowel space paradox. A sociophonetic study on Dutch," Ph.D. thesis, Radboud University, Nijmegen, The Netherlands.

Van de Velde, H. (**1996**). "Variatie en verandering in het gesproken standaardnederlands (1935–1993)" ["Variation and change in spoken standard Dutch (1935–1993)"], Ph.D. dissertation, Radboud University, Nijmegen, The Netherlands.

Van de Velde, H., Gerritsen, M., and van Hout, R. (**1996**). "The devoicing of fricatives in standard Dutch: A real-time study based on radio recording," Lang. Var. Change **8**(2), 149–175.

van Rij, J., Wieling, M., Baayen, R., and van Rijn, H. (**2022**). "itsadug: Interpreting time series and autocorrelated data using GAMMs (version 2.4.1) [computer software]," https://cran.r-project.org/web/packages/itsadug/ (Last viewed March 8, 2023).

van Son, R. (**2001**). "Protocol voor het oplijnen van fonetische transcripties met spraak" ("Protocol for aligning phonetic transcriptions to speech"), available at http://www.fon.hum.uva.nl/IFA-SpokenLanguageCorpora/IFAcorpus/SLcorpus/LabelProtocol/LabelProtocol.pdf (Last viewed May 31, 2021).

Vanvik, A. (**1972**). "A phonetic-phonemic analysis of Standard Eastern Norwegian," Norw. J. Linguist. **26**, 119–164.

Veenker, T. J. G. (**2012**). "The Zep experiment control application [computer software]," https://www.beexy.org/zep/ (Last viewed March 8, 2023).

Whalen, D. H., Gick, B., Kumada, M., and Honda, K. (**1999**). "Cricothyroid activity in high and low vowels: Exploring the automaticity of intrinsic F0," J. Phon. **27**, 125–142.

Wieling, M. (**2018**). "Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English," J. Phon. **70**, 86–116.

Wissing, D. (**1991**). "Regressiewe stemassimilasie in Afrikaans en Nederlands" ("Regressive voice assimilation in Afrikaans and Dutch"), S. Afr. J. Linguist. Suppl. 11, 132–156.

Wood, S. N. (**2017**). *Generalized Additive Models: An Introduction with R*, 2nd ed. (Chapman and Hall/CRC, Boca Raton, FL).