v1.12.9000 (2019/05/23)

Longer vowel duration correlates with greater tongue root displacement: Acoustic and articulatory data from Italian and Polish

Stefano Coretta

Abstract

Voiced stops tend to be preceded by longer vowels and produced with a more advanced tongue root than voiceless stops. The duration of a vowel is modulated by the voicing of the stop that follows and in many languages vowels are longer when followed by voiced stops. Tongue root advancement is known to be an articulatory mechanism which ensures the right pressure conditions for the maintenance of voicing during closure as dictated by the Aerodynamic Voicing Constraint. In this paper, it is argued that vowel duration and tongue root advancement enter in a direct statistical relation. Drawing from acoustic and ultrasound tongue imaging data from 17 speakers of Italian and Polish, it is shown that tongue root advancement is initiated during the vowel, and that vowel duration and tongue root position at vowel offset are positively correlated. Longer vowel durations correspond to greater tongue root advancement. It is further proposed that the later closure onset of voiced stops within a temporally stable interval is responsible for both greater root advancement and shorter closure durations in the context of voiced stops.

1 Introduction

It is well known that voiced stops are almost universally characterised by two phonetic correlates: advanced tongue root and increased duration of the preceding vowel (Westbury 1983; Lisker 1974; Fowler 1992). While a lot of work has been done on each of these aspects separately, less is known about their relation. In this paper, I propose a link between the position of the tongue root at the onset of a post-vocalic stop and the duration of the vowel preceding that stop. In an exploratory study of the articulatory correlates of stop voicing, it was found that tongue root advancement—a mechanism known to facilitate voicing during stop closure—is initiated during the production of the vowel preceding the stop. This replicates previous work on tongue root position. Furthermore, the results of this study indicate that the acoustic duration of the vowel is positively correlated with tongue root position, such that longer vowel durations correspond to greater tongue root advancement. Such correlation is shown to derive from the timing of the consonantal closure relative to the preceding vowel.

1.1 Tongue root position and voicing

One of the differences in supra-glottal articulation between voiced and voiceless stops concerns the position of the tongue root relative to the front-back dimension of the oral tract. It has been repeatedly observed that the tongue root is in a more front position in voiced stops compared to voiceless stops (Kent & Moll 1969; Perkell 1969; Westbury 1983). This has been attributed to the fact that the initiation and maintenance of vocal fold vibration (i.e. voicing) requires a difference in air pressure between the cavities below and above the glottis. Specifically, the sub-glottal pressure needs to be higher than the supra-glottal pressure. In other words, there must be a positive trans-glottal air pressure differential (van den Berg 1958; Rothenberg 1967). This property of voicing is formally known as the Aerodynamic Voicing Constraint (Ohala 2011). When the oral tract is completely occluded during the production of a stop closure, the supra-glottal pressure quickly increases, due to the incoming airstream from the lungs. Such pressure increase can hinder the ability to sustain vocal fold vibration during closure, to the point voicing ceases.

An articulatory solution to counterbalance the increased pressure is to enlarge the supra-glottal cavity by advancing the root of the tongue. In the context of articulatory adjustments, a distinction between passive and active gestures is generally drawn (see for example Rothenberg 1967). A passive enlargement of the oral cavity is the product of the incoming airflow, the pressure of which expands the pliable soft tissues of the cavity walls. On the other hand, active expansion is achieved by muscular activity, which can in turn be purposive (produced with the goal of cavity expansion) or non-purposive. While Rothenberg (1967) recognises that the distinction between purposive and non-purposive active gestures can be at times blurry, it is nonetheless important to note that the qualification of a gesture as active does not automatically implies a speaker's intention to produce the obtained result.

Rothenberg (1967) further calculates that the walls of the oral tract can absorb the incoming airflow for 20 to 30 ms by passive expansion, after which the sub- and supra-glottal pressures would equalise and voicing cease. Based on these estimates, a passive expansion of the pharyngeal walls is thus not generally sufficient to maintain voicing during the closure of a stop. Reaching a complete ballistic forward gesture would require the tongue root about 70 to 90 ms (Rothenberg 1967). Given that voiced stop closures are on average shorter than that (the mean duration is about 64 ms in Luce & Charles-Luce 1985), it is expected that the movement could be initiated during the production of the vowel, so that an appreciable amount of advancement is obtained when closure is achieved. Furthermore, Westbury (1983) finds that tongue root advancement is initiated before the achievement of full closure and that there is a forward movement even in some cases of voiceless stops, although the rate and magnitude of the advancement are consistently higher in voiced stops. Finally, tongue root adjustments seem to target more specifically lingual consonants, while the tongue body is more involved in labials (Perkell 1969; Westbury 1983).

However, the relation between tongue root advancement and voicing is a complex one. First, tongue root advancement is not the only mechanism for sustaining voicing during a stop (Rothenberg 1967; Westbury 1983; Ohala 2011) and it has a certain degree of idiosyncrasy (Ahn 2018). For example, a cross-lingusitically common difference between voiceless and voiced stops concerns their respective closure durations. The closure of voiced stops is generally longer than that of voiceless stops (Lisker 1957; Umeda 1977; Van Summers 1987; Davis & Van Summers 1989; de Jong 1991). A shorter closure favours maintenance of vocal fold vibration by ensuring that the pressure build-up in the oral cavity does not equalise the sub-glottal and supra-glottal pressures (at which point voicing would stop). Other solutions which can help sustaining voicing during closure include larynx lowering (Riordan 1980), slackening of the vocal folds Halle & Stevens (1967), opening of the velopharyngeal port (Yanagihara & Hyde 1966), and producing a retroflex occlusion (Sprouse et al. 2008). Moreover, (Ahn 2018) finds that not all the speakers she surveyed did show tongue root advancement, and a few had rather the reverse pattern.

Second, implementation of tongue root advancement can be decoupled from the actual presence of vocal fold vibration. In Westbury (1983), advancement of the tongue root is found in some productions of voiceless stops. This is counterintuitive, since tongue root advancement is generally considered to be a feature of voiced stops which require voicing-related pressure adjustments. Moreover, Ahn (2015, 2018) and Ahn & Davidson (2016) looked at utterance-initial stops and found that the tongue root is more advanced in the phonologically voiced stops independent of whether they are implemented with vocal fold vibration or not.

To summarise, tongue root advancement is a common articulatory solution employed to counterbalance the increase in supra-glottal pressure and maintaining voicing during the production of at least lingual voiced stops. While this gesture is not exclusive of voiced stops and it can be implemented even in the absence of vocal fold vibration, tongue root advancement seems to be a robust correlate of voicing.

1.2 Vowel duration and voicing

The results discussed here are part of a larger study which focusses on the effect of consonant voicing on preceding vowel durations. A great number of studies showed that, cross-linguistically, vowels tend to be longer when followed by voiced obstruents than when they are followed by voiceless ones (House & Fairbanks 1953; Peterson & Lehiste 1960; Chen 1970; Klatt 1973; Lisker 1974; Farnetani & Kori 1986; Fowler 1992; Hussein 1994; Esposito 2002; Lampp & Reklis 2004; Durvasula & Luo 2012). This so-called 'voicing effect' has been reported in a variety of languages, including (but not limited to) English, German, Hindi, Russian, Arabic, Korean, Italian, and Polish (see Maddieson & Gandour 1976 and Beguš 2017 for a more comprehensive list).

Italian and Polish offer an opportunity to study of the articulatory aspects of the voicing effect, given their reported differences in magnitude/presence of the effect and the relative ease of comparison. While Italian has been consistently reported as a voicing-effect language (Caldognetto et al. 1979; Farnetani & Kori 1986; Esposito 2002), some studies found an effect in Polish (Slowiaczek & Dinnsen 1985; Nowak 2006; Malisz & Klessa 2008; Coretta 2018) while others did not (Keating 1984; Jassem & Richter 1989).

Coretta (2018) argues, based on the acoustics of the same data reported here, that the stressed vowels of disyllabic (CV́CV) words in Italian and Polish are 16 ms longer (SE = 4.4) when followed by a voiced stop. The high degree of intra-speaker variation, backed up by statistical modelling, also indicates that these languages possibly behave similarly in regards to the voicing effect. Finally, the temporal distance between two consecutive stop releases in CV́CV words is not affected by the voicing of the second consonant. The duration of the release to release interval is stable across voicing contexts. Within this interval, the timing of the onset of the stop closure produces differences in the respective durations of vowel and closure, following a mechanism of temporal compensation (Lindblom 1967; Slis & Cohen 1969b,a; Lehiste 1970b,a). A later closure onset results in a long vowel and a short closure, while an earlier closure onset corresponds to a short vowel and a long closure. Since the closure of voiceless stops is longer than that of voiced stops, it follows that vowels are shorter when followed by the former than when followed by the latter.

1.3 This study

Previous research has established that tongue root advancement and longer vowel durations are two common correlates of voicing. In particular, voicing during closure can be maintained by advancing the tongue root during the production of voiced stops (which is possibly initiated earlier than the closure onset) and that vowels followed by voiced stops tend to be longer than vowels followed by voiceless stops. The acoustic data further revealed that the duration of the stop closure bears on the duration of the preceding vowel, by means of a kind of compensatory mechanism.

The results from the articulatory data of this study, which will be discussed in the following sections, offer new insights on the link between closure and vowel duration. We will see that the relative timing of the closure also modulates the degree of tongue root advancement found at closure onset, thus creating a three-way network of relations with vowel duration and tongue root position. More specifically, the timing of the closure onset within the release-to-release interval decides the duration of the vowel, the duration of the closure, and the degree of tongue root advancement. Finally, it will be argued that a later closure onset as in the case of voiced stops has the double advantage of producing both a short closure duration and greater tongue root advancement, features both known to comply with the Aerodynamic Voicing Constraint.

2 Methodology

2.1 Participants

Participants were recruited in Manchester (UK), and Verbania (Italy) Eleven native speakers of Italian (5 females, 6 males) and 6 native speakers of Polish (3 females, 3 males) participated in this study. Most speakers of Italian are originally from the North of Italy, while 3 are from Central Italy. The Polish speakers came from different parts of Poland (2 from the west, 3 from the centre, and 1 from the east). This study has been approved by the School of Arts, Languages, and Culture Ethics committee of the University of Manchester (REF 2016-0099-76). The participants signed a written consent and received a monetary compensation of £10.

2.2 Equipment

Simultaneous recordings of audio and ultrasound tongue imaging were obtained in the Phonetics Laboratory at the University of Manchester (UK) or in a quiet room in Verbania (Italy). An Articulate Instruments Ltd™ system was used for this study.

The system is made of a TELEMED Echo Blaster 128 unit, an Articulate Instruments Ltd™ P-Stretch synchronisation unit, and a FocusRight Scarlett Solo pre-amplifier. A TELEMED C3.5/20/128Z-3 ultrasonic transducer (20mm radius, 2-4 MHz) and a Movo LV4-O2 Lavalier microphone were used respectively for the acquisition of ultrasonic and audio data. The ultrasonic probe was placed in contact with the sub-mental triangle, aligned with the mid-sagittal plane. A metallic headset designed by Articulate Instruments Ltd™ (2008) was used to hold the probe in a fixed position and inclination relative to the head. The acquisition of the mid-sagittal ultrasonic and audio signals was achieved with the software Articulate Assistant Advanced (AAA, v2.17.2) running on a Hawlett-Packard ProBook 6750b laptop with Microsoft Windows 7. The synchronisation of the ultrasonic and audio signals was performed by AAA after recording by means of a synchronisation signal produced by the P-Stretch unit. The ranges of the ultrasonic settings were: 43-68 frames per second, 88-114 number of scan lines, 980-988 pixel per scan line, field of view 71-93°, pixel offset 109-263, depth 75-180 mm. The audio signal was sampled at 22050 Hz (16-bit).

2.3 Materials

Disyllabic words of the form $C_1V_1C_2V_2$ were used as targets, where C_1 = /p/, V_1 = /a, o, u/, C_2 = /t, d, k, g/, and V_2 = V_1 (e.g. *pata*, *pada*, *poto*, etc.), giving a total of 12 target words, used both for Italian and Polish.[1] The resulting words are nonce words, with a few exceptions, and they were presented in the languages' respective writing conventions (see Appendix A). A labial stop was chosen as the first consonant to reduce possible coarticulation with the following vowel.[2] Central/back vowels only were included in the target words for two reasons. First, high and mid front vowels tend to be difficult to image with ultrasound, given their greater distance from the ultrasonic probe when compared with back vowels. Second, high and mid front vowels usually produce less tongue displacement from and to a stop consonant. This characteristic can make it more difficult to identify gestural landmarks using the methodology discussed in Section 2.5. Since the focus of the study was to explore differences in the closing gesture of voiceless and voiced stops, only lingual consonants have been included (the closure of labial stops cannot of course be imaged with ultrasound). The sentence *Dico X lentamente* 'I say X slowly' in Italian, and *Mówię X teraz* 'I say X now' for Polish functioned

[1]Note that stressed vowels in open syllables in Italian are long (Renwick & Ladd 2016). Moreover, /o/ is used here for typographical simplicity to indicate the mid-back vowels of Italian and Polish, although they do differ in quality. See Krämer (2009), Renwick & Ladd (2016), and Gussmann (2007).

[2]However, note that Westbury (1983) and Vazquez-Alvarez & Hewlett (2007) report tongue body lowering in the context of labial stops.

as frames for the test words. Speakers were instructed to read the sentences without pauses and to speak at a comfortable pace.

2.4 Procedure

The participants familiarised themselves with the sentence stimuli at the beginning of the session. Headset and probe were then fitted on the participant's head. The participant read the sentence stimuli, which were presented on the computer screen in a random order, while the audio and ultrasonic signals were acquired simultaneously. The random list of sentences was read 6 times consecutively (with the exception of IT02, who repeated the sentences 5 times only). Due to software constraints, the order of the sentences within participant was kept the same for each of the six repetitions. The participant could optionally take breaks between one repetition and the other. Sentences with hesitations or speech errors were immediately discarded and re-recorded. A total of 1212 tokens (792 from Italian, 420 from Polish) were obtained.

2.5 Data processing and statistical analysis

The audio data was subject to force alignment using the SPeech Phonetisation Alignment and Syllabification software (SPPAS, Bigi 2015). The outcome of the automatic alignment was then manually corrected, according to the recommendations in Machač & Skarnitzl (2009). The onset and offset of V1 in the $C_1V_1C_2V_2$ test words were respectively placed in correspondence of the appearance and disappearance of higher formant structure in the spectrogram. Vowel duration was calculated as the duration of the V1 onset to V1 offset interval. Speech rate was measured as the number of syllables in the sentence (8 in Italian and 6 in Polish) divided by the duration of the sentence in seconds.

The displacement of the tongue root was obtained from the ultrasonic data according to the procedure used in Kirkham & Nance (2017). Smoothing splines were automatically fitted to the visible tongue contours in AAA. Manual correction was then applied in cases of clear tracking errors. A fan-like frame consisting of 42 equidistant radial lines superimposed on the ultrasonic image was used as the coordinate system. The origin of the 42 fan-lines coincides with the (virtual) origin of the ultrasonic beams, such that each fan-line is parallel to the direction of the ultrasonic scan lines. Tongue root displacement was thus calculated as the displacement of the fitted spline along a selected vector (Strycharczuk & Scobbie 2015), see Figure 1. For each participant, the fan-line with the highest standard deviation of displacement within the area corresponding to the speaker's tongue root was chosen as the tongue root displacement

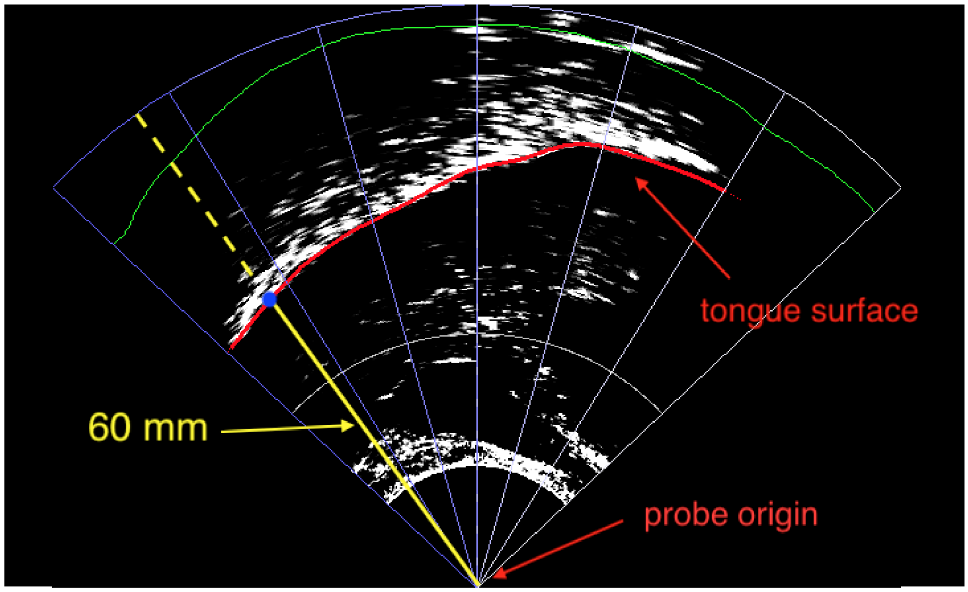

Figure 1: Schematics of the operationalisation of tongue root position, based on Kirkham & Nance (2017). The tongue root surface corresponds to the lower edge of the white band in the image. The tongue tip is on the right side. The outline of the fan-like coordinate systems is shown. The yellow line starting from the probe origin is the selected fan-line from which tongue root position is calculated (see text for the method of fan-line selection). Tongue root position thus corresponds to the distance (in millimetres) between the probe origin and the intersecting point of the tongue surface with the selected fan-line.

vector. A Savitzky–Golay smoothing filter (second-order, frame length 75 ms) was applied to the raw displacement. Displacement values for analysis are taken from the smoothed displacement signal. Tongue root displacement was obtained from a static time point (the onset of the closure of C2) and along the duration of the vowel. The displacement values along the vowel duration were extracted at time points corresponding to real ultrasonic video frames. Given the average frame rate is 55 frames per second, values are sampled about every 20 ms.

Statistical analysis was performed in R v3.5.2 (R Core Team 2018). Linear mixed-effects models were fitted with lme4 v1.1-19 (Bates et al. 2015). Factor terms were coded with treatment contrasts (the reference level is the first listed for each factor): C2 voicing (voiceless, voiced), vowel (/a/, /o/, /u/). Speech rate was centred for inclusion in the statistical models, by subtracting the mean speech rate across all speakers from the calculated speech rate values. Centring ensures the intercepts are interpretable. *t*-tests with Satterthwaite's approximation to degrees of freedom on the individual terms were used to obtain *p*-values using lmerTest v3.0-1 (Kuznetsova et al. 2017; Luke 2017). An effect is considered significant if the *p*-value is below the alpha level ($\alpha = 0.05$). Generalised additive mixed models were fitted with mgcv v1.8-26 (Wood 2011, 2017). The smooths used thin plate regression splines as basis (Wood 2003). The ordered factor difference smooths method described in Sóskuthy (2017) and Wieling (2018) was used to model the effect of factor terms in GAMs. The models were fitted by maximum likelihood (ML) and autoregression in the residuals was controlled with a first-order autoregressive model.

Significance testing of the relevant predictors was achieved by comparing the ML score of the full model with the score of a null model (in which the relevant predictor is dropped), using the `compareML()` function of the itsadug package (van Rij et al. 2017). A preliminary analysis indicated that including either language or C2 place of articulation as predictors produced respective *p*-values above the alpha level, without affecting the estimates of the other terms. Section 4.3 further discusses the idiosyncratic behaviour of the tongue root observed between speakers, which does not seem to pattern in any way with their native language. For these reasons, these variables were not included in the models reported here and will not be discussed. Future research is warranted to ascertain language-related differences and possible effects of place of articulation.

2.6 Open Science statement

Following recent practices which encourage scientific transparency (Crüwell et al. 2018; Berez-Kroeker et al. 2018; Roettger 2019), data and analysis code are available on the Open Science Framework (Coretta 2019).

3 Results

3.1 Tongue root position at C2 closure onset

Figure 2 shows raw data points and boxplots of the position of the tongue root at C2 closure onset when C2 is voiceless (left) and voiced (right). Since the position of the tongue root in millimetres depends on the speaker's anatomy and on the probe location, scaled tongue root position is used in this plot (note though that the unscaled data is used in statistical modelling). As a trend, the position of the tongue root is more advanced if C2 is voiced compared to its position when C2 is voiceless.

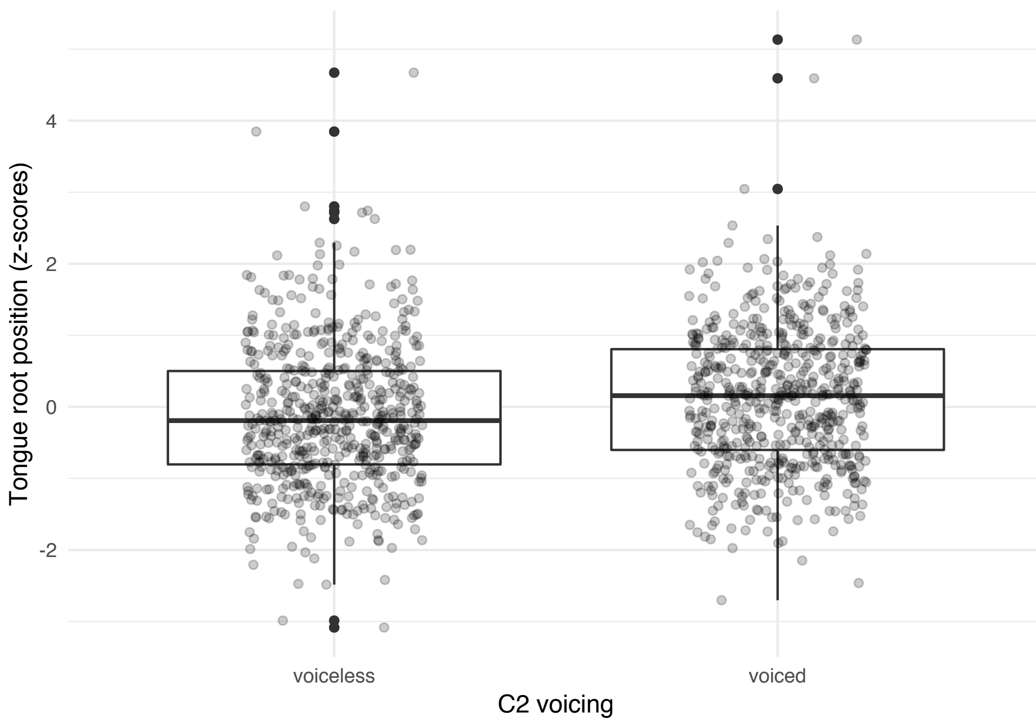

Figure 2: Raw data and boxplots of tongue root position in voiceless and voiced stops at closure onset. Higher values indicate advancement.

A linear mixed-effects model with tongue root position as the outcome variable was fitted with the following predictors (Table 2): fixed effects for C2 voicing (voiceless, voiced), centred speech rate (as number of syllables per second, centred), vowel (/a/, /o/, /u/); by-speaker and by-word random intercepts (a by-speaker random coefficient for C2 voicing led to singular fit, so it was not included in the final model). The effects of C2 voicing and vowel are significant according to *t*-tests with Satterthwaite's approximation to degrees of freedom. The tongue root at C2 closure onset is 0.77 mm (SE = 0.35) more front when C2 is voiced, and it is 1.87 mm (SE = 0.42) more retracted if V1 is /o/.

3.2 Tongue root position during V1

The position of the tongue root during the articulation of V1 was assessed with generalised additive mixed models (GAMM). A GAMM was fitted to tongue root position with the following terms (Table 3): C2 voicing as a parametric term; a smooth term over centred speech rate, a smooth term over V1 proportion with a by-C2 voicing difference smooth, a tensor product interaction over V1 proportion and centred speech rate; a factor random smooth over V1 proportion by speaker (penalty order = 1). A chi-squared test on the ML scores of the full model and a model excluding C2 voicing indicates that C2 voicing significantly improves fit ($\chi(3) = 7.758$, $p = 0.001$). Figure 3 shows that the root advances during the production of the vowel, relative to its position at V1 onset. This forward movement is observed both in the context of a following voiced stop and in that of a following voiceless stop. However, the magnitude of the movement is greater in the former. At V1 offset (= C2 closure onset), the graph suggests a difference in tongue root position of about 1 mm.

3.3 Correlation between tongue root position and V1 duration

A second linear mixed regression was fitted to tongue root position to assess the effect of V1 duration on root position (Table 4). The following terms were included: centred V1 duration (in milliseconds), centred speech rate (as number of syllables per second), vowel (/a/, /o/, /u/), C2 place of articulation (coronal, velar); an interaction between centred V1 duration and vowel; by-speaker and by-word random intercept (a by-speaker random coefficient for V1 duration led to non-convergence, so it was not included in the final model). All predictors and the V1 duration/vowel interaction are significant. V1 duration and tongue root position are positively correlated: The longer the vowel, the more advanced the tongue root is at V1 offset ($\hat{\beta} = 0.065$ mm, SE = 0.007). The effect is stronger with /a/ than with /o/ and /u/ (see Figure 4).

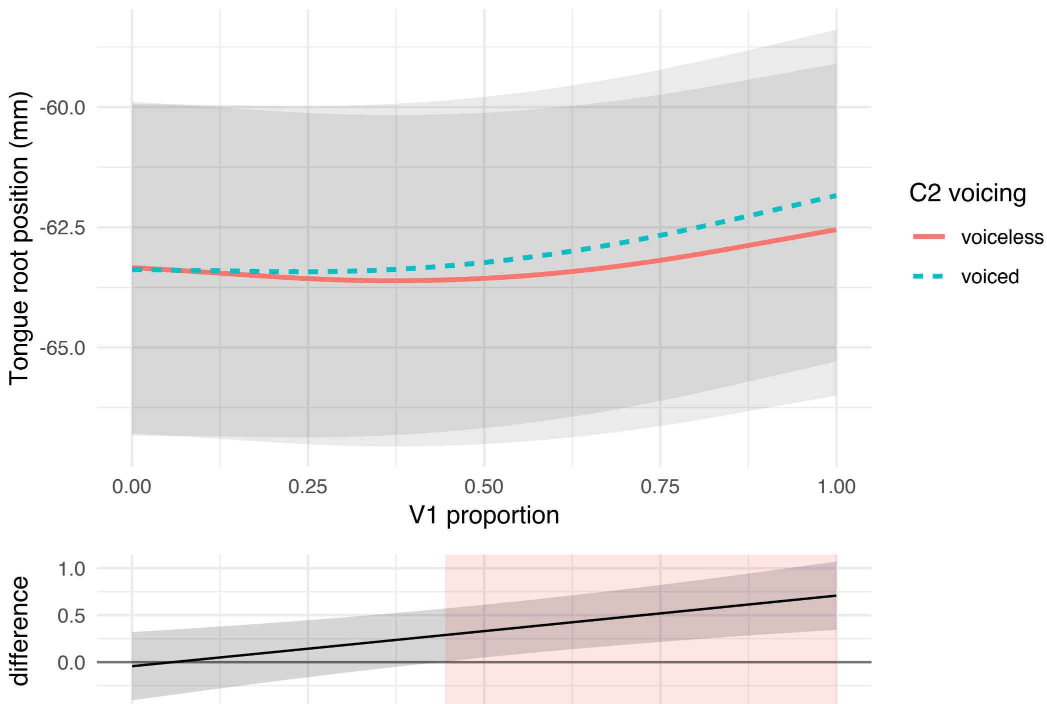

Figure 3: Predicted tongue root position (top figure) during vowels preceding voiceless and voiced stops, with 95% confidence intervals, and difference smooth (bottom figure). Higher values of tongue root position indicate a more advanced root. The shaded red area in the difference smooth indicates where the two curves are different. Predictions from a GAMM (see Section 3.2).

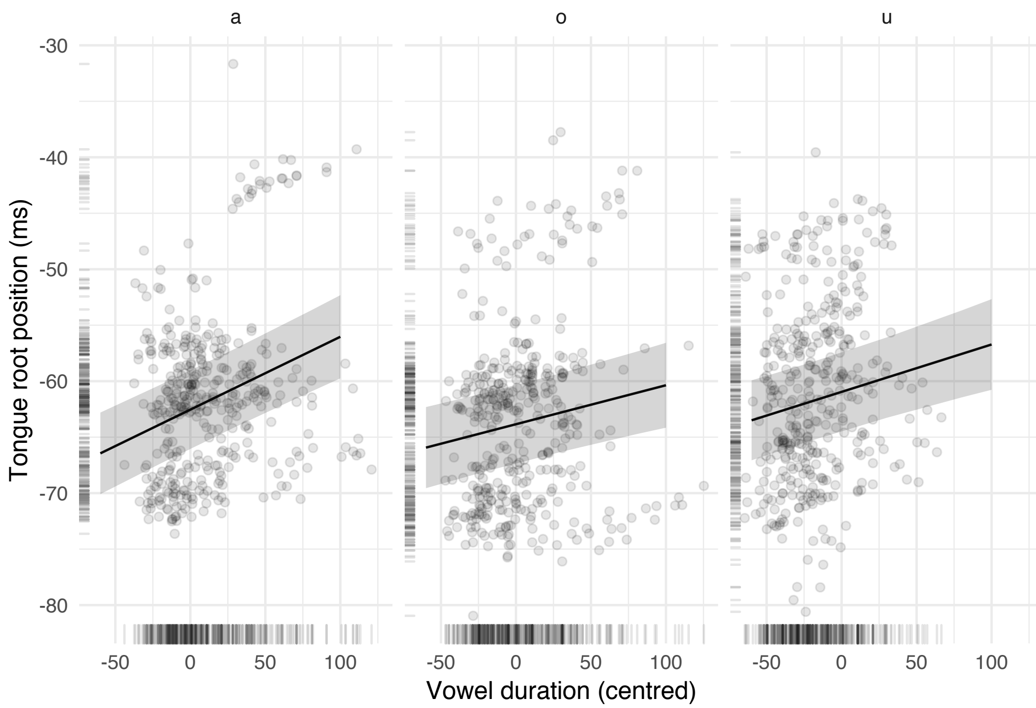

Figure 4: Raw data, regression lines, and 95% confidence intervals of the correlation between vowel duration and tongue root position for each vowel (/a/, /o/, and /u/). The regression line and confidence intervals are from a mixed-effects model (see Section 3.3).

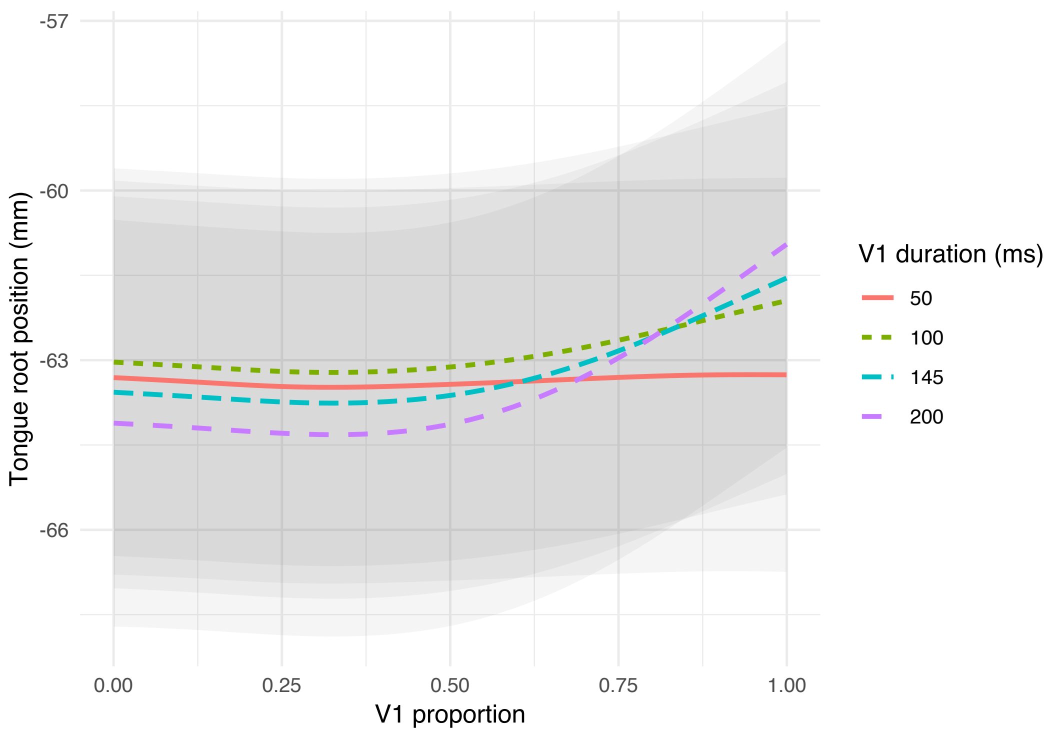

Figure 5: Predicted tongue root position during vowels at 4 exemplifying values of vowel duration, with 95% confidence intervals. Predictions from a GAMM (see Section 3.4).

3.4 Tongue root position during V1 as a function of V1 duration

The effect of V1 duration on tongue root position during V1 was modelled by fitting a GAMM with the following terms (Table 5): tongue root position as the outcome variable, smooth terms over V1 duration and V1 proportion, a tensor product interaction over V1 proportion and V1 duration; a factor random smooth over V1 proportion by speaker (penalty order = 1). The full model with the tensor product interaction over V1 proportion and V1 duration has better fit according to model comparison with a model without the interaction ($\chi(3) = 12.559$, $p < 0.001$). Figure 5 shows the estimated root trajectories at four values of vowel duration. The general trend is that the forward movement of the root during the vowel is greater the longer the duration of the vowel (Figure 5). Moreover, the trajectory curvature increases with vowel duration: Shorter vowels have a flatter trajectory of tongue root advancement.

4 Discussion

4.1 Voicing, tongue root position and vowel duration

The results of this study of voicing and vowel duration in Italian and Polish revealed a few patterns in the relation between consonant voicing, tongue root position, and vowel duration. Unsurprisingly, the position of the tongue root at vowel offset is more front when the following stop is voiced than when the following stop is voiceless in both surveyed languages. This finding aligns with the results of previous work on English (Kent & Moll 1969; Perkell 1969; Westbury 1983; Ahn 2018). When looking at the position of the tongue root during the vowel, it was found that the root starts advancing during the articulation of the vowel. Westbury (1983) found the same pattern in English. Moreover, similarly to the results in Westbury (1983), some tongue root advancement during the production of the vowel is found even when C2 is voiceless.

A possible reason for the presence of such a small degree of advancement in voiceless lingual stops is offered by arguments in relation to the absence of advancement in labials (voiced or voiceless). Westbury (1983) proposes that the articulation of the closure of lingual stops mechanically involves movements of the tongue root, so that, in order to keep a constant oral cavity volume, the root moves forward while the tongue body moves upward. On the other hand, the tongue can move freely in labial stops since their closure involves the lips. This idea is supported by the 'trough effect' (Vazquez-Alvarez & Hewlett 2007), i.e. VCV sequences involving a labial stop show tongue body lowering, and by the fact that voiced labials tend to resort to tongue body lowering rather than tongue root advancement as a mechanism for voicing maintenance (Perkell 1969; Westbury 1983; Ahn 2018). The small degree of advancement in voiceless lingual stops could then as well be a mechanic consequence of the tongue moving upward for producing the stop closure.

The data discussed here also suggest that tongue root position is positively correlated with vowel duration, such that longer vowels show a more advanced tongue root at vowel offset (= closure onset) than shorter vowels. Said correlation exists independent of the voicing status of the consonant following the vowel (compatible with the finding that even voiceless stops have some degree of advancement). The correlation between tongue root and vowel duration could be interpreted as to indicate that the onset of the forward gesture of the root is timed relative to a landmark preceding the closure, independent of the duration of the vowel. The timing of the stop closure along the advancement movement would sanction the degree of advancement found at closure onset.

The dynamic data of tongue root advancement during the articulation of the vowel indicates that vowels followed by voiced stops have greater tongue root advancement at vowel offset than vowels followed by voiceless stops, in accordance with the results from the static analysis at vowel offset. Moreover, a significant interaction between vowel duration and the trajectory shape was found. Shorter vowels have a flatter trajectory, while the curvature of the trajectory in longer vowels has a greater curvature.

When comparing the effects of vowel duration and speech rate on tongue root position, though, we are faced with a paradox. Both variables have a positive effect on tongue root position, so that longer vowels and higher speech rates imply a more advanced root. However, speech rate has a negative effect on vowel duration (and segments duration in general), such that higher speech rates are correlated with shorter vowel durations (this holds for this data). If higher speech rates mean shorter vowels and shorter vowels imply a less advanced root, we should also find less advancement with higher speech rates. However, the results indicate the opposite, and higher speech rates are correlated with more root advancement. However, a regression model on the position of the tongue root at *vowel onset* suggests that speech rate is positively correlated with tongue root position at vowel onset. The tongue root is already in a more advanced position at vowel onset when the speech rate is high, so that, if vowel duration is held constant, more advancement is expected at vowel offset with higher speech rates even when higher speech rate has a negative effect on vowel duration.

The articulatory patterns observed in this paper contribute to the understanding of the acoustic patterns discussed in previous work. If we take the release of the consonant preceding the vowel as a reference point, a delayed consonant closure could ensure that, by the time closure is made, an appreciable amount of tongue root advancement is achieved. Other things being equal, an increase in cavity volume increases the time required to reach trans-glottal pressure equalisation, which would cause cessation of voicing. This mechanism thus contributes to the maintenance of voicing during the stop closure.

The closure of voiced stops is achieved later (relative to the preceding consonant release) compared to the closure of voiceless stops. Moreover, the temporal distance between the releases of the two consecutive stops in CV́CV words is not affected by the voicing category of the second stop. Given the stability of the release to release interval duration, the delay in producing a full closure seen in the context of voiced stops has thus a double advantage: (1) A greater degree of tongue root advancement is achieved at vowel offset/closure onset, and (2) the stop closure is shorter. Both of these articulatory features are compliant with the requirements dictated by the aerodynamic voicing constraint. A more advanced tongue root ensures that the trans-glottal pressure differential is sufficient for voicing to be sustained, and a shorter closure reduces the pressure build-up during the stop closure. To conclude, it is proposed that the combined action of a temporally stable release to release interval and the differential timing of the closure onset of voiceless vs. voiced stops contribute to both the acoustic patterns of vowel and closure duration and the articulatory patterns of tongue root position.

4.2 Estimates of tongue root displacement

It is worth to briefly discuss the estimated difference in tongue root position between voiceless and voiced stops and its significance. The estimated magnitude of such difference is 0.77 mm (SE = 0.35). The 95% confidence interval for the difference is approximately within the range 0-1.5 mm. Rothenberg (1967) argues that the anterior wall of the lower pharynx (corresponding to the tongue root) can move by 5 mm along the antero-posterior axis. Figure 1 in Kirkham & Nance (2017) suggests that the tongue root of one of the Twi speakers recorded is about 4 mm more front in /e/ (a +ATR vowel) than in /ɛ/ (a -ATR vowel). Given that a difference of 4 mm in root position can produce a substantially distinct acoustic output in vowels (like the two different phonemes of Twi), it makes sense to expect that differences in tongue root position as driven by consonantal factors should be of some magnitude smaller, like the once found in this study. Moreover, the data presented here indicates that for every millisecond increase in vowel duration there is a 0.065 mm increase in tongue root advancement (see Section 3.3). If a maximal ballistic forward movement of the tongue root takes between 70 and 90 ms (Rothenberg 1967), we can calculate the maximum displacement plausible to be between 4.55 to 5.85 mm (0.065 mm times 70–90 ms). These values are in agreement with the maximum root displacement of 5 mm estimated by Rothenberg.

The results of this study also shed some light on timing aspects of tongue root advancement. As mentioned in the previous section, the correlation between tongue root position and vowel duration could be a consequence of the timing of the advancement gesture. In order to obtain such correlation, the onset of the gesture (during the articulation of the vowel) should be at a fixed distance from an earlier reference point (like the vowel onset or the preceding consonant offset) such that the timing of consonant closure will create the correlation seen in the data. Although ideally the timing of the onset of the advancing gesture should be fixed, the velocity of the gesture itself could be different depending on the voicing of the following consonant. It is possible that the velocity will be greater in the context of voiced stops, especially if the advancing gesture in this context is executed with greater muscular force. Unfortunately, a preliminary screening of the current data was inconclusive as to whether timing and velocity are similar or differ in the voiceless and voiced contexts, due to the difficulty in identifying the onset of the advancing gesture. Further data should be collected with the aim of testing the hypothesis that the timing of the gesture onset is the same in

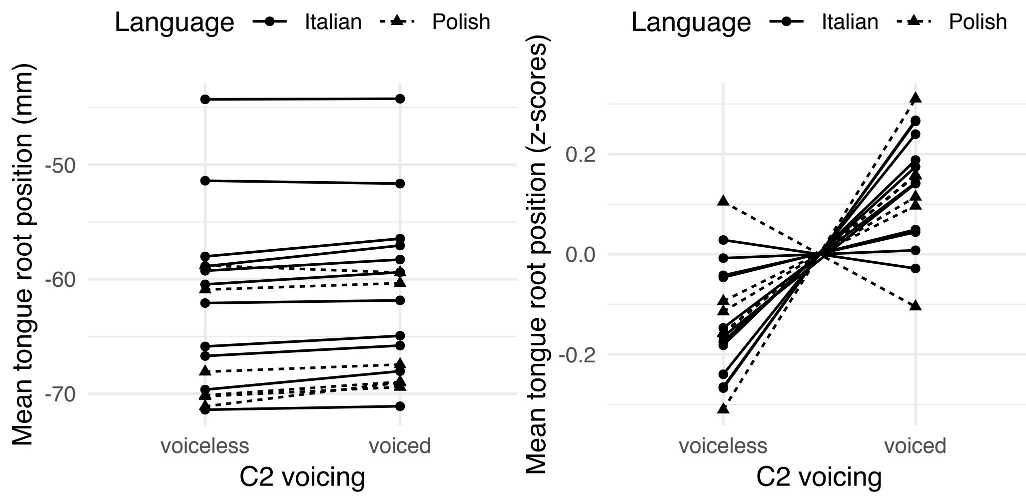

Figure 6: Slope plots of mean torngue root position in voiceless and voiced stops at closure onset, by-speaker. The plot on the left has raw position values in millimetres, while the plot on the right shows standardised values (z-scores) by speaker. See text for details.

voiceless and voiced contexts, while the velocity of the gesture should differ.

Although the results of this study are in agreement with previous work, the correlation between tongue root position and vowel duration needs to be replicated by expanding the enquired contexts to other types of consonants and vowels, and with other languages. Investigating the relative phasing of tongue root and body gestures in lingual and labial consonants is also necessary to clarify the mechanisms that could underlie the gestural timing of stop closure and tongue root advancement. Moreover, while the paper so far has focussed on group-level trends, it should be noted that, as found in other studies on the tongue root, individual speakers show a somewhat high degree of variability. The following section discusses this point.

4.3 Individual differences

The results presented in Section 3 and discussed in Section 4 are group-level patterns of the population sampled in the present study. However, the data is characterised by a certain degree of individual-level differences. Figure 6 shows two slope plots of mean tongue root position depending on C2 voicing for each speaker. In each plot, the two means of each speaker are linked by a line that shows the difference (or lack thereof) in means. Solid lines are Italian speakers, while dashed lines are Polish speakers. The *y*-

axis of the left plot is the raw mean position in millimetres, while that of the right plot is the standardised values (z-scores) of the mean position. An upward-slanted slope line indicates that the mean tongue root position in the voiced condition is higher, while a downward-slanted slope is interpreted as a decrease in mean root position. A flat slope suggests there is no difference in means between the voiceless and voiced condition.

This plot show that all three possibilities of slope direction are found in the data. The mean value of tongue root position of a voiced C2 relative to that of a voiceless stop is greater in some speakers, smaller in others, and similar in yet other speakers. Moreover, no discernible pattern can be found between speakers of Italian and Polish. Speakers of both languages show more or less the same range of variation. However, as we have seen in Section 3, the estimated overall effect of C2 voicing is robust and it implies a more advanced tongue root in voiced stops. The right plot of Figure 6 confirms this point visually. Two speakers show a declining slope (one is Italian and the other Polish), one speaker has a virtually flat slope, while all the others have an increasing slope at varying degrees. Note that the individual variation across speakers found in this data is qualitatively comparable to that in Ahn (2018).

The mean difference in tongue root position in voiceless vs. voiced stops has been calculated for each speaker from the raw data. Figure 7 plots the speakers' mean differences, with the respective standard error bars. The bottom 7 speakers (3 Polish, 4 Italian) show either a weak negative difference (the tongue root is slightly more advanced in voiceless stops) or a weak positive difference with wide standard errors which include 0. The remaining 11 speakers have a more robust positive difference (the tongue root is more advanced in voiced stops). Finally, speakers of each language do not cluster together, reiterating the observation made above that language does not seem to be an informative parameter.

Finally, interesting individual patterns can also be seen in the trajectories of tongue root position. Figure 8 shows these trajectories for all the speakers (note that the y-axis of each plot is on a different scale, so magnitude comparisons should not be made visually). Speakers IT01, IT03, and PL04 in particular have a somewhat categorical distinction in tongue root position during vowels followed by voiceless vs. voiced stops. Such tongue root distinction is implemented across the total duration of the vowel, rather than towards the end (as suggested by the results from the aggregated data, see Section 3.2). The phonological literature reports cases in which the difference in tongue root position in vowels is enhanced, leading to phonological alternations or diachronic loss of the voicing distinction with maintenance of the tongue root distinction (see Vaux 1996 and references therein). The ultrasound data from this study offers articula-

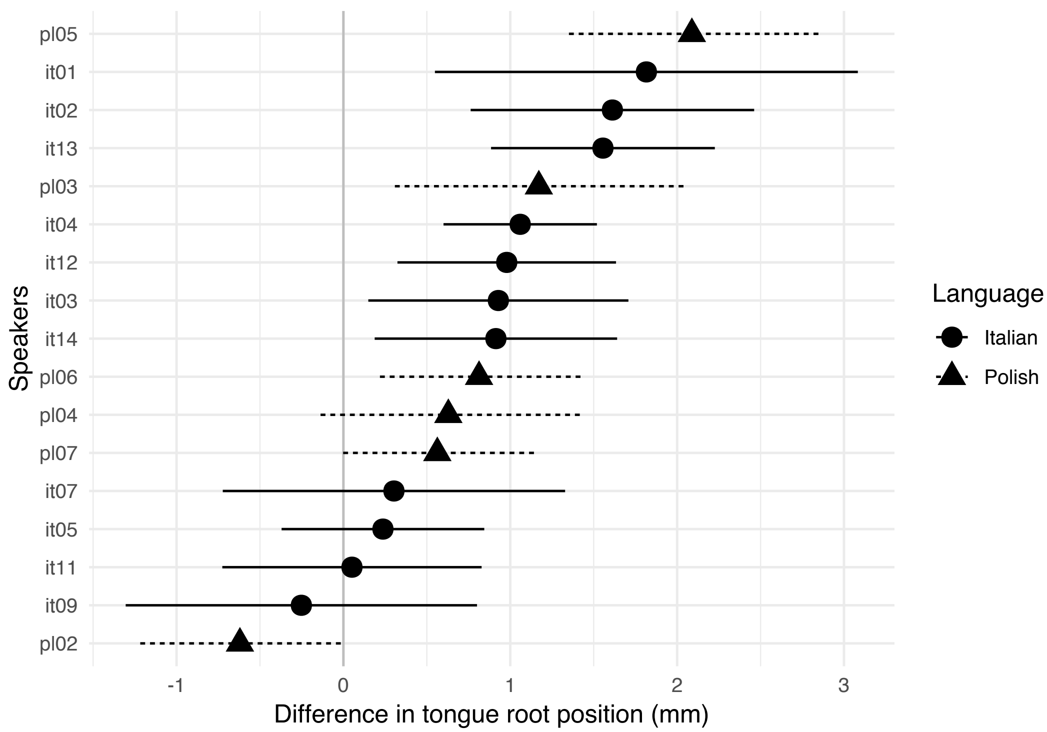

Figure 7: By-speaker raw mean difference in tongue root position between voiceless and voiced stops at closure onset (in millimetres). The horizontal segments are the standard errors of the mean differences.

tory evidence for a possible precursor of said phonological patterns.[3]

5 Conclusion

The maintenance of voicing during the closure of stops can achieved through a variety of articulatory mechanisms. Among these, shorter closure durations and cavity expansion by tongue root advancement are commonly observed solutions. Another robust correlate of consonant voicing is longer preceding vowel duration. This paper discussed articulatory data from an exploratory study of the effect of voicing on vowel duration first introduced in Coretta (2018). Similarly to what previously found for English, the tongue root at stop closure onset is more advanced in voiced than in voiceless stops in Italian and Polish. The average difference in tongue root position is 0.77 mm (SE = 0.35). By modelling the trajectory of the tongue root during the production of vowels preceding stops, it was found that the root starts advancing during the vowel, both preceding voiceless and voiced stops. The magnitude of the advancing gesture was however greater in the voiced context. Moreover, tongue root position and vowel duration were found to be positively correlated. Longer vowel durations correspond to greater tongue root advancement.

It was argued that the combined action of two factors contribute to the patterns observed: (1) The duration of the interval between two consecutive releases, and (2) the timing of the C2 closure onset within such interval. The release to release interval duration has been found not to be affected by the voicing of the second consonant. The later closure onset of voiced stops within the release to release interval (compared to voiceless stops) has the double advantage of producing a shorter closure duration and ensuring that enough tongue root advancement is reached by the time the stop closure is achieved. Both of these aspects comply with the aerodynamic voicing constraint (Ohala 2011) by delaying trans-glottal pressure equalisation (which would prevent vocal fold vibration). Future studies will need to test whether these findings replicate in Italian and Polish, and if they extend to other languages and contexts. In particular, further work on the relative differences in timing and velocity of the closing gesture and the root advancement gesture will be necessary to obtain a more in-depth understanding of the relation between consonant voicing, tongue root position, and vowel duration.

[3] All the examples in Vaux (1996) are on vowels *following* voiceless vs. voiced stops, rather than preceding, as in the current study. While beyond the scope of this paper, whether this is a systematic gap or not and how this relates to the present findings should be examined in future work.

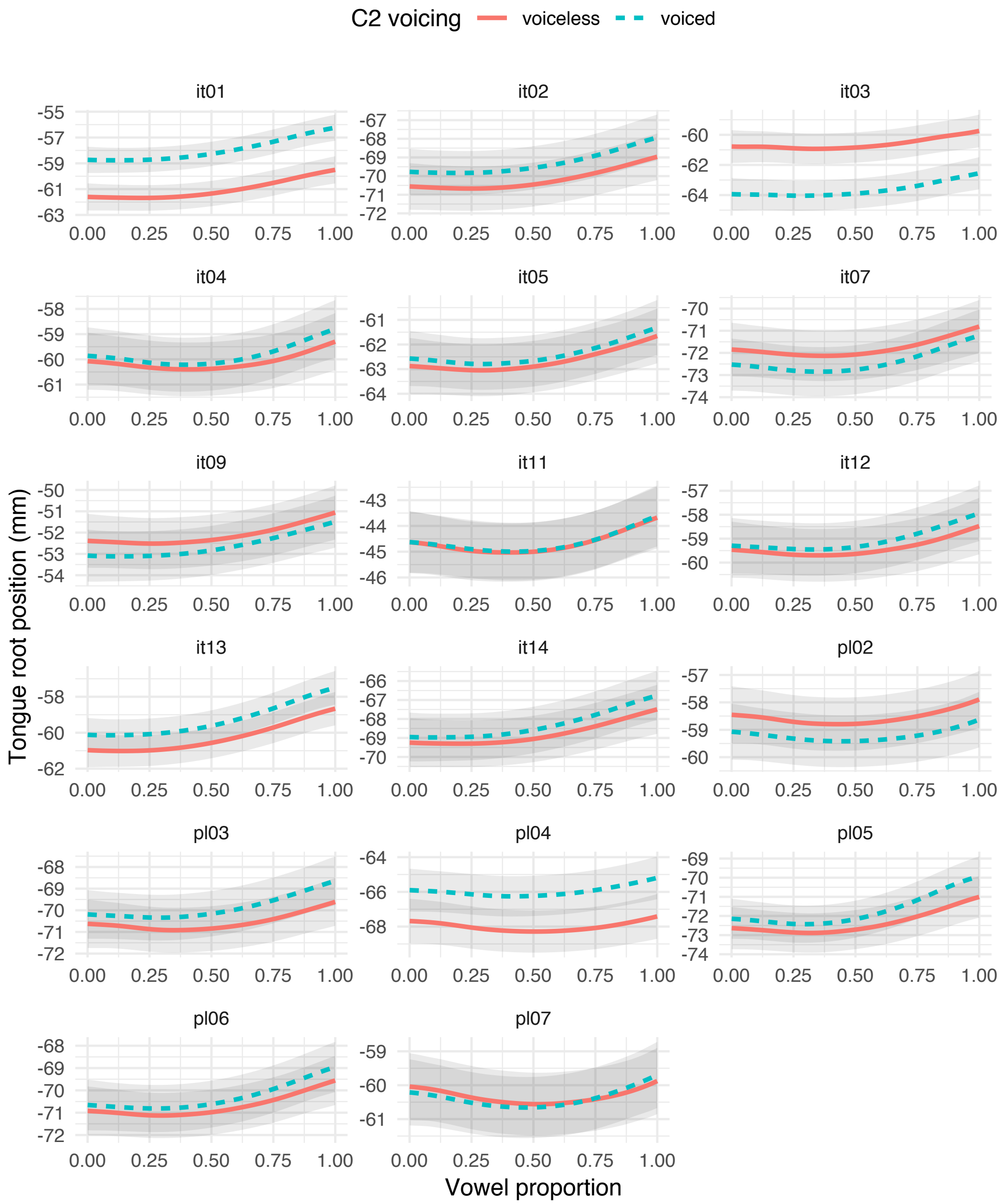

Figure 8: Predicted tongue root position during vowels followed by voiceless and voiced stops for each speaker. Predicted from a GAMM (see text). Note the different scales on the y-axis.

Table 1: The list of Italian and Polish target words. An asterisk indicates a real word.

Italian			Polish		
pata	poto*	putu	pata	poto	putu
pada	podo	pudu	pada*	podo	pudu
paca*	poco*	pucu	paka*	poko	puku
paga*	pogo	pugu	paga	pogo	pugu

Table 2: Summary of the linear mixed-effects model fitted to tongue root position at vowel offset (see Section 3.1)

Predictor	Estimate	SE	CI low	CI up	df	t-value	p-value	< α
Intercept	-62.1396	1.8113	-65.6898	-58.5895	17.1188	-34.3058	0.0000	*
Voicing = voiced	0.7689	0.3473	0.0881	1.4497	19.3947	2.2137	0.0390	*
Speech rate (centr.)	0.4114	0.2793	-0.1360	0.9588	1168.1100	1.4732	0.1410	
Vowel = /o/	-1.8742	0.4249	-2.7069	-1.0414	19.2874	-4.4112	0.0003	*
Vowel = /u/	0.0865	0.4270	-0.7503	0.9233	19.6974	0.2027	0.8415	

A Target words

See Table 1.

B Output of statistical models

See Table 2, Table 3, Table 4, Table 5.

References

Ahn, Suzy. 2015. The role of the tongue root in phonation of American English stops. Paper presented at Ultrafest VII.

Ahn, Suzy. 2018. The role of tongue position in laryngeal contrasts: An ultrasound study of english and brazilian portuguese. *Journal of Phonetics* 71. 451–467.

Ahn, Suzy & Lisa Davidson. 2016. Tongue root positioning in English voiced obstruents: Effects of manner and vowel context. *The Journal of the Acoustical Society of America* 140(4). 3221–3221.

Table 3: Summary of the GAM model fitted to tongue root position during V1 (see Section 3.2)

Predictor	Estimate	SE	EDF	Ref.DF	Statistic	p-value	< α
Intercept	-63.3328	1.7562			-36.0623	0.0000	*
Voicing = voiced	0.3311	0.1432			2.3122	0.0208	*
s(Speech rate (centr.))			7.5311	8.5159	4.4779	0.0000	*
s(Proportion)			3.6915	4.3646	10.4427	0.0000	*
s(Proportion): voiced			1.0122	1.0236	10.0413	0.0015	*
ti(Proportion, Speech Rate (c.))			2.1335	2.7679	2.9005	0.0429	*
s(Proportion, Speaker)			62.2821	152.0000	57.3445	0.0000	*

Table 4: Summary of the linear mixed-effects model for testing the correlation between tongue root position and V1 duration (see Section 3.3)

Predictor	Estimate	SE	CI low	CI up	df	t-value	p-value	< α
Intercept	-62.5793	1.7818	-66.0716	-59.0870	17.0874	-35.1212	0.0000	*
V1 duration (centr.)	0.0651	0.0073	0.0507	0.0795	955.6436	8.8558	0.0000	*
Speech rate (centr.)	1.2412	0.2903	0.6722	1.8102	1169.6885	4.2755	0.0000	*
Vowel = /o/	-1.3031	0.4597	-2.2040	-0.4021	18.3761	-2.8348	0.0108	*
Vowel = /u/	1.5863	0.5049	0.5967	2.5759	25.8255	3.1419	0.0042	*
V1 duration × /o/	-0.0303	0.0079	-0.0457	-0.0149	736.2314	-3.8504	0.0001	*
V1 duration × /u/	-0.0227	0.0090	-0.0403	-0.0052	751.2493	-2.5345	0.0115	*

Articulate Instruments Ltd™. 2008. Ultrasound stabilisation headset users manual: Revision 1.4. Edinburgh, UK: Articulate Instruments Ltd.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. doi:10.18637/jss.v067.i01.

Beguš, Gašper. 2017. Effects of ejective stops on preceding vowel duration. *The Journal of the Acoustical Society of America* 142(4). 2168–2184. doi:10.1121/1.5007728.

Berez-Kroeker, Andrea L., Lauren Gawne, Susan Smythe Kung, Barbara F. Kelly, Tyler Heston, Gary Holton, Peter Pulsifer, David I. Beaver, Shobhana Chelliah & Stanley Dubinsky. 2018. Reproducible research in linguistics: A position statement on data citation and attribution in our field. *Linguistics* 56(1). 1–18. doi:10.1515/ling-2017-0032.

Table 5: Summary of the GAM model fitted to tongue root position during V1 as a function of V1 duration (see Section 3.4)

Predictor	Estimate	SE	EDF	Ref.DF	Statistic	p-value	< α
Intercept	-63.0628	1.7397			-36.2487	0	*
s(V1 duration)			8.2416	8.8458	7.7094	0	*
s(Proportion)			3.9640	4.7069	17.9941	0	*
ti(Proportion, V1 duration)			2.8557	3.3236	8.9785	0	*
s(Proportion, Speaker)			59.9588	152.0000	65.7394	0	*

van den Berg, Janwillem. 1958. Myoelastic-aerodynamic theory of voice production. *Journal of Speech and Hearing Research* 1(3). 227–244. doi:10.1044/jshr.0103.227.

Bigi, Brigitte. 2015. SPPAS - Multi-lingual approaches to the automatic annotation of speech. *The Phonetician* 111–112. 54–69.

Caldognetto, Emanuela Magno, Franco Ferrero, Kyriaki Vagges & Maria Bagno. 1979. Indici acustici e indici percettivi nel riconoscimento dei suoni linguistici (con applicazione alle consonanti occlusive dell'italiano). *Acta Phoniatrica Latina* 2. 219–246.

Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22(3). 129–159.

Coretta, Stefano. 2018. An exploratory study of voicing-related differences in vowel duration as compensatory temporal adjustment in Italian and Polish. Submitted.

Coretta, Stefano. 2019. Longer vowel duration correlates with greater tongue root displacement: Acoustic and articulatory data from Italian and Polish [Research compendium]. Open Science Framework, `https://osf.io/e6btr/`.

Crüwell, Sophia, Johnny van Doorn, Alexander Etz, Matthew Makel, Hannah Moshontz, Jesse Niebaum, Amy Orben, Sam Parsons & Michael Schulte-Mecklenbeck. 2018. 8 easy steps to open science: An annotated reading list. PsyArXiv. doi:10.31234/osf.io/cfzyx.

Davis, Stuart & W. Van Summers. 1989. Vowel length and closure duration in word-medial VC sequences. *Journal of Phonetics* 17. 339–353.

Durvasula, Karthik & Qian Luo. 2012. Voicing, aspiration, and vowel duration in Hindi. *Proceedings of Meetings on Acoustics* 18. 1–10. doi:10.1121/1.4895027.

Esposito, Anna. 2002. On vowel height and consonantal voicing effects: Data from Italian. *Phonetica* 59(4). 197–231. doi:10.1159/000068347.

Farnetani, Edda & Shiro Kori. 1986. Effects of syllable and word structure on segmental durations in spoken Italian. *Speech communication* 5(1). 17–34. doi:10.1016/0167-6393(86)90027-0.

Fowler, Carol A. 1992. Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here. *Journal of Phonetics* 20(1). 143–165.

Gussmann, Edmund. 2007. *The phonology of Polish*. Oxford University Press.

Halle, Morris & Kenneth Stevens. 1967. Mechanism of glottal vibration for vowels and consonants. *The Journal of the Acoustical Society of America* 41(6). 1613–1613. doi:10.1121/1.2143736.

House, Arthur S. & Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America* 25(1). 105–113. doi:10.1121/1.1906982.

Hussein, Lutfi. 1994. *Voicing-dependent vowel duration in Standard Arabic and its acquisition by adult American students*: The Ohio State University dissertation.

Jassem, Wiktor & Lutoslawa Richter. 1989. Neutralization of voicing in Polish obstruents. *Journal of Phonetics* 17(4). 317–325.

de Jong, Kenneth. 1991. An articulatory study of consonant-induced vowel duration changes in English. *Phonetica* 48(1). 1–17. doi:10.1121/1.2028316.

Keating, Patricia A. 1984. Universal phonetics and the organization of grammars. *UCLA Working Papers in Phonetics* 59.

Kent, Raymond D. & Kenneth L. Moll. 1969. Vocal-tract characteristics of the stop cognates. *Journal of the Acoustical Society of America* 46(6B). 1549–1555.

Kirkham, Sam & Claire Nance. 2017. An acoustic-articulatory study of bilingual vowel production: Advanced tongue root vowels in Twi and tense/lax vowels in Ghanaian english. *Journal of Phonetics* 62. 65–81. doi:10.1016/j.wocn.2017.03.004.

Klatt, Dennis H. 1973. Interaction between two factors that influence vowel duration. *The Journal of the Acoustical Society of America* 54(4). 1102–1104. doi:10.1121/1.1914322.

Krämer, Martin. 2009. *The phonology of Italian*. Oxford: Oxford University Press.

Kuznetsova, Alexandra, Per Bruun Brockhoff & Rune Haubo Bojesen Christensen. 2017. `lmerTest` package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13). doi:10.18637/jss.v082.i13.

Lampp, Claire & Heidi Reklis. 2004. Effects of coda voicing and aspiration on Hindi vowels. *The Journal of the Acoustical Society of America* 115(5). 2540–2540. doi:10.1121/1.4783577.

Lehiste, Ilse. 1970a. Temporal organization of higher-level linguistic units. *The Journal of the Acoustical Society of America* 48(1A). 111–111. doi:10.1121/1.1974906.

Lehiste, Ilse. 1970b. Temporal organization of spoken language. In *Working papers in linguistics*, vol. 4, 96–114. doi:10.1121/1.1974906.

Lindblom, Björn. 1967. Vowel duration and a model of lip mandible coordination. *Speech Transmission Laboratory Quarterly Progress Status Report* 4. 1–29.

Lisker, Leigh. 1957. Closure duration and the intervocalic voiced-voiceless distinction in English. *Language* 33(1). 42–49. doi:10.2307/410949.

Lisker, Leigh. 1974. On "explaining" vowel duration variation. In *Proceedings of the Linguistic Society of America*, 225–232.

Luce, Paul A & Jan Charles-Luce. 1985. Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production. *The Journal of the Acoustical Society of America* 78(6). 1949–1957. doi:10.1121/1.392651.

Luke, Steven G. 2017. Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods* 49(4). 1494–1502. doi:10.3758/s13428-016-0809-y.

Machač, Pavel & Radek Skarnitzl. 2009. *Principles of phonetic segmentation*. Epocha.

Maddieson, Ian & Jack Gandour. 1976. Vowel length before aspirated consonants. In *UCLA Working papers in Phonetics*, vol. 31, 46–52.

Malisz, Zofia & Katarzyna Klessa. 2008. A preliminary study of temporal adaptation in Polish VC groups. In *Proceedings of speech prosody*, 383–386.

Nowak, Pawel. 2006. *Vowel reduction in Polish*: University of California, Berkeley dissertation.

Ohala, John J. 2011. Accommodation to the aerodynamic voicing constraint and its phonological relevance. In *Proceedings of the 17th International Congress of Phonetic Sciences*, 64–67.

Perkell, Joseph S. 1969. *Physiology of speech production: Results and implication of quantitative cineradiographic study*. Cambridge, MA: MIT Press.

Peterson, Gordon E. & Ilse Lehiste. 1960. Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America* 32(6). 693–703. doi:10.1121/1.1908183.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. `https://www.R-project.org`.

Renwick, Margaret & Robert D. Ladd. 2016. Phonetic distinctiveness vs. lexical contrastiveness in non-robust phonemic contrasts. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 7(1). 1–29. doi:10.5334/labphon.17.

Riordan, Carol J. 1980. Larynx height during english stop consonants. *Journal of Phonetics* 8. 353–360.

Roettger, Timo B. 2019. Researcher degrees of freedom in phonetic sciences. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 10(1). 1–27. doi:10.5334/labphon.147.

Rothenberg, Martin. 1967. *The breath-stream dynamics of simple-released-plosive production*, vol. 6. Basel: Biblioteca Phonetica.

Slis, Iman H. & Antonie Cohen. 1969a. On the complex regulating the voiced-voiceless distinction II. *Language and speech* 12(3). 137–155. doi:10.1177/002383096901200301.

Slis, Iman Hans & Antonie Cohen. 1969b. On the complex regulating the voiced-voiceless distinction I. *Language and speech* 12(2). 80–102. doi:10.1177/002383096901200202.

Slowiaczek, Louisa M. & Daniel A. Dinnsen. 1985. On the neutralizing status of Polish word-final devoicing. *Journal of phonetics* 13(3). 325–341.

Sóskuthy, Márton. 2017. Generalised additive mixed models for dynamic analysis in linguistics: A practical introduction. arXiv.org preprint, arXiv:1703.05339.

Sprouse, Ronald L., Maria-Josep Solé & John J. Ohala. 2008. Oral cavity enlargement in retroflex stops. *Proceedings of the 8th International Seminar on Speech Production, Strasbourg* 425–428.

Strycharczuk, Patrycja & James M. Scobbie. 2015. Velocity measures in ultrasound data. Gestural timing of post-vocalic /l/ in English. In *Proceedings of the 18th International Congress of Phonetic Sciences*, 1–5.

Umeda, Noriko. 1977. Consonant duration in American English. *The Journal of the Acoustical Society of America* 61(3). 846–858. doi:10.1121/1.381374.

van Rij, Jacolien, Martijn Wieling, R. Harald Baayen & Hedderik van Rijn. 2017. itsadug: Interpreting time series and autocorrelated data using GAMMs. R package version 2.3.

Van Summers, W. 1987. Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *The Journal of the Acoustical Society of America* 82(3). 847–863. doi:10.1121/1.395284.

Vaux, Bert. 1996. The status of ATR in feature geometry. *Linguistic Inquiry* 27(1). 175–182.

Vazquez-Alvarez, Yolanda & Nigel Hewlett. 2007. The 'trough effect': an ultrasound study. *Phonetica* 64(2-3). 105–121. doi:10.1159/000107912.

Westbury, John R. 1983. Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *The Journal of the Acoustical Society of America* 73(4). 1322–1336.

Wieling, Martijn. 2018. Analyzing dynamic phonetic data using generalized additive mixed modeling: a tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics* 70. 86–116. doi:10.1016/j.wocn.2018.03.002.

Wood, Simon. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)* 73(1). 3–36.

Wood, Simon. 2017. *Generalized additive models: An introduction with R*. Chapman and Hall/CRC 2nd edn.

Wood, Simon N. 2003. Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65(1). 95–114.

Yanagihara, Naoaki & Charlene Hyde. 1966. An aerodynamic study of the articulatory mechanism in the production of bilabial stop consonants. *Studia Phonologica* 4. 70–80.