

# An exploratory study of voicing-related differences in vowel duration as compensatory temporal adjustment in Italian and Polish

Stefano Coretta  
*The University of Manchester*  
Oxford Road, Manchester, M13 9PL,  
United Kingdom  
[stefano.coretta@manchester.ac.uk](mailto:stefano.coretta@manchester.ac.uk)

<sup>1</sup> **Keywords:** vowel duration, closure duration, voicing effect, compensation, Italian,  
<sup>2</sup> Polish

<sup>3</sup> **Abstract** Over a century of phonetic research has established the cross-  
<sup>4</sup> linguistic existence of the so called ‘voicing effect’, by which vowels tend to  
<sup>5</sup> be shorter when followed by voiceless stops and longer when the following  
<sup>6</sup> stop is voiced. However, no agreement is found among scholars regarding  
<sup>7</sup> the source of this effect, and several causal accounts have been advanced.  
<sup>8</sup> A notable one is the compensatory temporal adjustment account, according  
<sup>9</sup> to which the duration of the vowel is inversely correlated with the stop clo-  
<sup>10</sup> sure duration (voiceless stops having longer closure durations than voiced  
<sup>11</sup> stops). The compensatory account has been criticised due to lack of empiri-  
<sup>12</sup> cal support and its vagueness regarding the temporal interval within which  
<sup>13</sup> compensation is implemented. The results from an exploratory study of  
<sup>14</sup> Italian and Polish suggest that the duration of the interval between two  
<sup>15</sup> consecutive stop releases in CVCV words in these languages is not affected  
<sup>16</sup> by the voicing of the second stop. The durational difference of the first  
<sup>17</sup> vowel and the stop closure would then follow from differences in timing of  
<sup>18</sup> the VC boundary within this interval. While other aspects, like production  
<sup>19</sup> mechanisms related to laryngeal features effects and perceptual biases can-  
<sup>20</sup> not be ruled out, the data discussed here are compatible with a production  
<sup>21</sup> account based on compensatory mechanisms.

## 1 Introduction

Almost a hundred years of research have consistently shown that consonantal voicing has an effect on preceding vowel duration: vowels followed by voiced obstruents are longer than when followed by voiceless ones (Meyer 1904; Heffner 1937; House & Fairbanks 1953; Belasco 1953; Peterson & Lehiste 1960; Halle & Stevens 1967; Chen 1970; Klatt 1973; Lisker 1974; Laeufer 1992; Fowler 1992; Hussein 1994; Lampp & Reklis 2004; Warren & Jacks 2005; Durvasula & Luo 2012). This so called ‘voicing effect’ has been found in a considerable variety of languages.<sup>1</sup> These include (but are not limited to) English, German, French, Spanish, Hindi, Russian, Italian, Arabic, and Korean (see Maddieson & Gandour 1976 for a more comprehensive, but still not exhaustive list).<sup>2</sup> Despite of the plethora of evidence in support of the *existence* of the voicing effect, agreement hasn’t been reached regarding its *source*.

Several proposals have been put forward in relation to the possible source of the voicing effect (see Sóskuthy 2013 and Beguš 2017 for an overview). Some of the proposed mechanisms for the emergence of the voicing effect refer to properties of speech production. A notable production account, which will be the focus of this study, is based on compensatory temporal adjustments (Lindblom 1967; Slis & Cohen 1969a; b; Lehiste 1970b; a). According to this account, the voicing effect follows from the reorganisation of gestures within a unit of speech the duration of which is not affected by stop voicing. The duration of such a unit is held constant across voicing contexts, while the duration of voiceless and voiced obstruents differs. The closure of voiceless stops is longer than that of voiced stops (Lisker 1957; Van Summers 1987; Davis & Van Summers 1989; de Jong 1991). As a consequence, vowels followed by voiceless stops (which have a long closure) are shorter than vowels followed by voiced stops (which have a short closure). Advocates of a compensatory mechanism propose two prosodic units as the scope of the temporal adjustment: the syllable (and, equivalently, the VC sequence or vowel-to-vowel interval, Lindblom 1967; Farnetani & Kori 1986), and the word (Slis & Cohen 1969a; b; Lehiste 1970b; a). How-

<sup>1</sup> One of the first attestations of the term ‘voicing effect’ can be attributed to Mitleb (1982). Another term used to refer to the same phenomenon is ‘pre-fortis clipping’, probably introduced by Wells (1990).

<sup>2</sup> A typological note. Most languages reported having a voicing effect come from the Indo-European family. Others are from a pool of widely studied languages. It is thus of vital importance that future studies look at other language families and underdocumented/underdescribed languages.

54 ever, the compensatory temporal adjustment account has been criticised in  
55 subsequent work.

56 Empirical evidence and logic challenge the proposal that the syllable or  
57 the word have a constant duration and hence drive compensation. First,  
58 Lindblom's 1967 argument that the duration of the syllable is constant is  
59 not supported by the findings in Chen (1970) and Jacewicz et al. (2009).  
60 Chen (1970) rejects a syllable-based compensatory mechanism in the light  
61 of the fact that the duration of the syllable is affected by consonant voicing.  
62 Jacewicz et al. (2009) further show that the duration of monosyllabic words  
63 in American English changes depending on the voicing of the coda conso-  
64 nant. Second, although the results in Slis & Cohen (1969b) suggest that the  
65 duration of disyllabic words in Dutch is constant whether the second stop  
66 is voiceless or voiced, it does not follow from this fact that compensation  
67 should necessarily target the vowel preceding the stop. Indeed, it is logi-  
68 cally possible that the following unstressed vowel could be the target of the  
69 compensation, therefore differences in preceding vowel duration still call  
70 for an explanation.

71 The compensatory temporal adjustment account has been further chal-  
72 lenged on the basis of the so called 'aspiration effect' (Maddieson & Gan-  
73 dour 1976), by which vowels are longer when followed by aspirated stops  
74 than when followed by unaspirated stops. In Hindi, vowels before voiceless  
75 unaspirated stops are short, vowels followed by voiced aspirated stops are  
76 long, and vowels followed by voiced unaspirated and voiceless aspirated  
77 stops are in between and have similar durations. Maddieson & Gandour  
78 (1976) find no compensatory pattern between vowel and consonant dura-  
79 tion. The consonant /t/, which has the shortest duration, is preceded by  
80 the shortest vowel, and vowels before /d/ and /t<sup>h</sup>/ have the same duration  
81 although the durations of the two consonants are different. Maddieson &  
82 Gandour (1976) argue that a compensatory explanation for differences in  
83 vowel duration cannot be maintained.

84 However, a re-evaluation of the way consonant duration is measured in  
85 Maddieson & Gandour (1976) might actually turn their findings in favour of  
86 a compensatory account. Due to difficulties in detecting the release of the  
87 consonant of interest, consonant duration in Maddieson & Gandour (1976)  
88 is measured from the closure of the relevant consonant to the release of the  
89 following, (e.g., in *ab sāth kaho*, the duration of /t<sup>h</sup>/ in *sāth* is calculated as  
90 the interval between the closure of /t<sup>h</sup>/ and the release of /k/). This measure  
91 includes the burst and aspiration (if present) of the consonant following  
92 the target vowel. Slis & Cohen (1969a), however, state that the inverse  
93 relation between vowel duration and the following consonant applies to

closure duration, and not to the entire *consonant* duration.<sup>3</sup> If an inverse relation exists between vowel and closure duration, the inclusion of burst and/or aspiration clearly alters this relationship.

Indeed, the study on Hindi voicing and aspiration effects conducted by Durvasula & Luo (2012) indicates that closure duration, measured from closure onset to closure offset, decreases according to the hierarchy voiceless unaspirated > voiced unaspirated > voiceless aspirated > voiced aspirated, which closely resembles the order of increasing vowel duration in Maddieson & Gandour (1976). Nonetheless, Durvasula & Luo (2012) do not find a negative correlation between vowel duration and consonant closure duration, but rather a (small) *positive effect*. Vowel duration increases with closure duration when voicing and aspiration are taken into account. However, as noted in Beguš (2017), it is likely that this result is a consequence of not controlling for speech rate. A small negative effect of closure duration can turn positive if the effect of speech rate (which is positive) is greater, given the cumulative nature of these effects.

de Jong (1991) finds partial support for a compensatory mechanism between vowel and closure duration in an electro-magneto-articulometric study of two American English speakers. The duration of vowels in nuclear accented, pre-, and post-nuclear accented position is weakly negatively correlated with closure duration (the slope coefficients range between -0.12 and -0.35, meaning that the amount of durational compensation is between 10% and 35%). While the magnitude of the correlation is too weak to univocally support compensation, the direction of the correlation is correct (i.e. a negative correlation).

Further evidence for a compensatory account and a negative correlation between vowel and closure duration comes from the effect of a third type of consonants, namely ejectives. Beguš (2017) finds that in Georgian (which contrasts aspirated, voiced, and ejective consonants) vowels are short when followed by voiceless aspirated stops, longer before ejective stops, and longest when followed by voiced stops. Crucially, stop closure duration follows the reversed pattern: closure is short in voiced stops, longer in ejectives, and longest in voiceless aspirated stops. Moreover, vowel duration is inversely correlated with closure across the three phonation types. Beguš (2017) mentions the possibility that the negative correlation is an artefact of the vowel and closure intervals sharing a boundary. This annotation bias could generate negative correlations (by which the vowel would

<sup>3</sup> In this paper, I use the term *relation* to mean a categorical pattern of entailment (like in ‘a long vowel entails a short closure’), while the term *correlation* is reserved to a statistical correlation of two continuous variables.

131 shorten and the closure would lengthen by the same amount when, for ex-  
132 ample, the boundary is placed to the left of the ‘actual’ boundary). How-  
133 ever, Beguš shows with a cross-annotator analysis that this was not the case.  
134 Beguš (2017) argues that these findings support temporal compensation (al-  
135 though not univocally, see Beguš 2017: Section V, and Section 4.2 of this  
136 paper).

137 To summarise, a mechanism of compensatory temporal adjustment has  
138 been proposed as the pathway to the emergence of the voicing effect. Ac-  
139 cording to such an account, the difference in vowel duration before conso-  
140 nants varying in voicing (and possibly other phonation types) is the outcome  
141 of a compensation between vowel and closure duration. After reviewing the  
142 critiques advanced by Chen (1970) and Maddieson & Gandour (1976), and  
143 in face of the results in Slis & Cohen (1969b), de Jong (1991) and Beguš  
144 (2017), a temporal compensation mechanism gains credibility. However,  
145 issues about the actual implementation of the compensation mechanism still  
146 remain. While compensatory temporal adjustments are plausible in light of  
147 the reviewed literature, we are still left with the necessity of identifying a  
148 speech interval the duration of which is not affected by the voicing of the  
149 post-vocalic consonant, and within which compensation can be logically  
150 implemented.

## 151 1.1 The present study

152 This paper reports on selected results from a broader exploratory study that  
153 investigates the relationship between vowel duration and consonant voicing  
154 from both an acoustic and articulatory perspective. Synchronised record-  
155 ings of audio, ultrasound tongue imaging, and electroglottography were car-  
156 ried out to enable a data-driven approach to the analysis of features related  
157 to the voicing effect in the context of disyllabic (CVCV) words in Italian and  
158 Polish.<sup>4</sup> This study was not designed to test the compensatory account, but  
159 rather to collect synchronised articulatory and acoustic data on the voic-  
160 ing effect. Moreover, the design of the study has been constrained by the  
161 use of ultrasound articulatory techniques (see Section 2). Since the tongue  
162 imaging and electroglottographic data don’t bear on the main argument put  
163 forward here, only the results from acoustics will be discussed.

---

<sup>4</sup> As per Cysouw & Good (2013), the glossonyms *Italian* and *Polish* as used here to refer, respectively, to the languoids Italian [GLOTTOCODE: ital1282] and Polish [GLOTTOCODE: poli1260].

Italian and Polish reportedly differ in the magnitude (or presence) of the effect of stop voicing on vowel duration. On the other hand, the typical realisation of phonological voiced stops in these languages are similar (but see [Huszthy 2016](#) and [Schwartz & Arndt 2018](#) for a phonological and phonetic discussion on laryngeal aspects of Italian and Polish respectively).<sup>5</sup> [Cyran \(2011\)](#) argues for a distinction between voicing and aspirating varieties of Polish, based on phonological arguments. [Waniek-Klimczak \(2011\)](#), on the other hand, cautiously argues that a possible change in progress in Polish is affecting the VOT values of voiceless stops in pre-stressed position.

The non-clear status of Polish laryngeal phonology/phonetics could be seen as a hindrance affecting the comparison with Italian. However, based on data from Italian, [Kirby & Ladd](#) propose that the distinction between voicing and aspirating languages itself ([Beckman et al. 2013](#)) cannot be straightforwardly mapped onto phonetics, and they remind us that ‘the production of laryngeal contrasts of all kinds are considerably more complex’ than generally described in the phonological literature ([Kirby & Ladd 2016: 2409](#)). Since this study focusses on the effect of post-stressed stops on preceding vowel durations, we believe that the comparison between Italian and Polish is still feasible, even in the case Polish voiceless pre-stressed stops are articulated with longer VOT values. Given that Italian and Polish share some features of the segmental and prosodic make-up of their phonological systems, the design of the experimental material and comparison of the results were facilitated. For these reasons, these languages offer an opportunity to investigate differences that could reveal mechanisms underlying the voicing effect, at least on a general level.

Italian has been unanimously reported as a voicing-effect language ([Magno Caldognetto et al. 1979](#); [Farnetani & Kori 1986](#); [Esposito 2002](#)). The mean difference in vowel duration when followed by voiceless vs. voiced consonants ranges between 22 and 24 ms in these studies, with longer vowels followed by voiced consonants. The mean differences are based on 3 speakers in [Farnetani & Kori 1986](#) and 7 speakers in [Esposito 2002](#). [Magno Caldognetto et al. \(1979\)](#) don’t report estimates of vowel duration, just the direction of the effect, but the study is based on 10 speakers.

The results regarding the presence and magnitude of the effect in Polish are instead mixed. [Ślowiaczek & Dinnsen \(1985\)](#) find that vowels followed by word-final underlyingly voiced stops are 10–15 ms longer in 5 Polish speakers, although [Jassem & Richter \(1989\)](#) did not replicate their results.

<sup>5</sup> Polish neutralises the voicing contrast word-finally, although the contrast is maintained word-medially ([Gussmann 2007](#)).



Similarly, [Keating \(1984\)](#) reports a difference of 2 ms in the duration of stressed vowels in disyllabic words from 24 speakers, which the author argues to be non-significant. On the other hand, [Nowak \(2006\)](#) finds that vowels followed by voiced stops are 4.5 ms longer in the 4 speakers recorded. [Malisz & Klessa \(2008\)](#) argue based on data from 40 speakers that the magnitude of the voicing effect in Polish is highly idiosyncratic, and claim that their results are inconclusive on this matter. While they do not report estimates from the 40 speakers, a table with mean vowel durations from 4 suggests a mean difference before voiceless vs. voiced stops of 3.5 ms. Finally, [Strycharczuk \(2012\)](#) reports a non-significant effect in 6 speakers in pre-sonorant word-final position.

The variety of results concerning the voicing effect in Polish could be related to differences in methodology. However, no clear pattern between studies which find a voicing effect and those which don't can be identified. For example, the studies reviewed here looked at either word-final or word-medial stops, controlled or read speech, speakers with a low or advanced proficiency in English. However, in all the individual cases both a positive and a negative result are reported depending on the study. What might be more relevant, though, is that the estimates of the difference in vowel duration are generally very low, between 3.5 and 15 ms. Given the small magnitude of the difference, it is likely that the failure to obtain significant *p*-values in some studies are due to low statistical power, rather than because of absence of the effect (as also hinted in [Beguš 2017](#), see arguments in [Roettger 2019](#) and [Nicenboim et al. 2018](#)).

The acoustic data from the study discussed here suggests that (1) a voicing effect can be detected both in Italian and Polish, and that (2) the duration of the interval between two consecutive stop releases (the release to release interval) is not affected by the voicing of the second consonant in both languages. This finding is compatible with a compensatory temporal adjustment account by which the timing of the closure onset of the stop following the vowel within said interval determines the respective durations of vowel and closure.

## 2 Method

### 2.1 Participants

Participants were sought in Manchester (UK), and in Verbania (Italy). Seventeen subjects in total participated in this study. Eleven subjects are native

speakers of Italian (5 female, 6 male), while six are native speakers of Polish (3 female, 3 male). The Italian speakers are from the North and Centre of Italy (8 speakers from Northern Italy, 3 from Central Italy). The Polish group has 2 speakers from Western Poland, 3 speakers from Central Poland, and 1 speaker from Eastern Poland. For more information on the sociolinguistic details of the speakers, see Appendix B. Ethical clearance for this study was obtained from the University of Manchester (REF 2016-0099-76). The participants signed a written consent and received a monetary compensation of £10.

## 2.2 Equipment

The acquisition of the audio signal was achieved with the software Articulate Assistant Advanced™ (AAA, v2.17.2, [Articulate Instruments Ltd™ 2011](#)) running on a Hewlett-Packard ProBook 6750b laptop with Microsoft Windows 7. Audio recordings were sampled at 22050 Hz (16-bit) and saved in a proprietary format (.aa0). A FocusRight Scarlett Solo pre-amplifier and a Movo LV4-O2 Lavalier microphone were used for audio recording. The microphone was placed at the level of the participant's mouth on one side, at a distance of about 10 cm. The microphone was clipped onto a metal headset worn by the participant, which was part of the ultrasonic equipment.

## 2.3 Materials

The target stimuli were disyllabic words with  $C_1V_1C_2V_2$  structure, 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.).<sup>6</sup> Most are nonce words, although inevitably some combinations produce real words both in Italian (4 words) and Polish (2 words, see Appendix C). The lexical stress of the target words was placed by speakers of both Italian and Polish on  $V_1$ , as intended.

The make-up of the target words was constrained by the design of the experiment, which included ultrasound tongue imaging (UTI). Front vowels

<sup>6</sup> Italian has both a mid-low [ɔ] and a mid-high [o] back vowel in its vowel inventory. These vowels are traditionally described as two distinct phonemes ([Krämer 2009](#)), although both their phonemic status and their phonetic substance are subject to a high degree of geographical and idiosyncratic variability ([Renwick & Ladd 2016](#)). As a rule of thumb, stressed open syllables in Italian (like the ones used in this study) have [ɔ:] (vowels in penultimate stressed open syllables are long) rather than [o:] ([Renwick & Ladd 2016](#)). On the other hand, Polish has only a mid-low back vowel phoneme /ɔ/ ([Gussmann 2007](#)). For the sake of typographical simplicity, the symbol /o/ will be used here for both languages.



are difficult to be imaged with UTI, since their articulation involves tongue surface positions which are particularly far from the ultrasonic probe, hence reducing the visibility of the tongue contour. For this reason, only central and back vowels were included. Since one of the variables of interest in the study was the closing gesture of C<sub>2</sub>, only lingual consonants were used. A labial stop was chosen as the first consonant to reduce possible coarticulation with the following vowel (although see [Vazquez-Alvarez & Hewlett 2007](#)). The number of target words was kept low to reduce the time required for completing the task, since the ultrasonic equipment can get very uncomfortable for the speaker when worn for more than 15/20 minutes.

The target words were embedded in a frame sentence. Controlling for meaning, segmental and prosodic make-up between languages proved to be difficult. The frames are *Dico X lentamente* ‘I say X slowly’ in Italian (following [Hajek & Stevens 2008](#)), and *Mówię X teraz* ‘I say X now’ in Polish. These sentences were chosen in order to maintain a similar intonation contour across languages.

## 2.4 Procedure

The participant was asked to read the sentences with the target words which were presented on the computer screen. The order of the sentences was randomised for each participant. Participants read the list of randomised sentence stimuli 6 times. Due to software constraints, the order of the list was kept the same across the six repetitions within each participant. The reading task lasted between 15 and 20 minutes, with optional short breaks between one repetition and the other. The total session time was around 45 minutes. Before the start of the experiment, the participants were spoken to in their mother tongue to try and reduce exposition to English prior to being recorded. Instructions were also given in their respective mother tongues. Each speaker read a total of 12 sentences for 6 times (with the exceptions of IT02, who repeated the 12 sentences 5 times), which yields a grand total of 1212 tokens (792 from Italian, 420 from Polish).

The experiment was carried out in two locations: in the sound attenuated booth of the Phonetics Laboratory at the University of Manchester, and in a quiet room in a field location in Italy (Verbania, Northern Italy). In both locations the equipment and procedures were the same. Data collection started in December 2016 and ended in March 2018.

**Table 1:** Criteria for the identification of acoustics landmarks.

landmark		criteria
vowel onset	(V1 onset)	Appearance of higher formants in the spectrogram following the release of /p/ (C1)
vowel offset	(V1 offset)	Disappearance of the higher formants in the spectrogram preceding the target consonant (C2)
consonant onset	(C2 onset)	Corresponds to V1 offset
closure onset	(C2 closure onset)	Corresponds to V1 offset
consonant offset	(C2 offset)	Appearance of higher formants of the vowel following C2 (V2); corresponds to V2 onset
consonant release	(C1/C2 release)	Automatic detection + manual correction ( <a href="#">Ananthapadmanabha et al. 2014</a> )

## 2.5 Data processing and measurements

The audio recordings were exported from AAA in the .wav format for further processing. The sample and bit rate were kept as upon recording (22050 Hz, 16-bit). A forced aligned transcription was accomplished through the SPeech Phonetisation Alignment and Syllabification software (SPPAS, [Bigi 2015](#)). The outcome of the automatic annotation was manually corrected for the relevant boundaries, according to the criteria in Table 1 based on [Machač & Skarnitzl \(2009\)](#). Segmentation boundaries not used in the analyses have not been checked to speed up processing. The releases of C1 and C2 were detected automatically by means of a Praat scripting implementation of the algorithm described in [Ananthapadmanabha et al. \(2014\)](#), and subsequently corrected if necessary. The identification of the stop release was not possible in 99 tokens (8%) of C1 and 265 tokens (22%) of C2 out of 1212. This was due either to the absence of a clear burst in the waveform and spectrogram, or the realisation of voiced stops as voiced fricatives. Most of the fricativised tokens come from three speakers of Central Italian, IT12,

IT13, and IT14, a variety of Italian known to show processes of lenition (Hualde & Nadeu 2011).

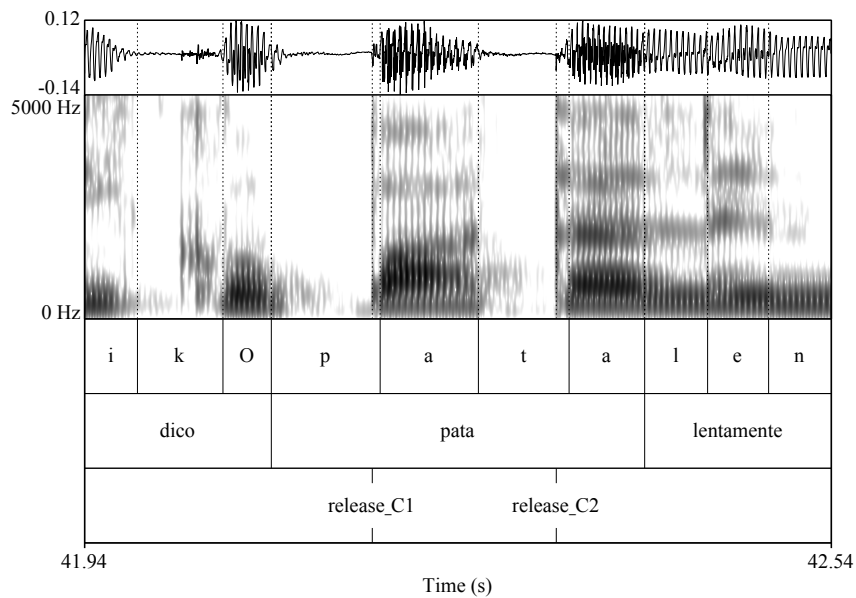
Moreover IT12 and IT14 produced several tokens of voiceless stops with voicing during closure (in some cases the closure was completely voiced). These tokens have been used in the analyses, because (1) the actual presence or absence of voicing during closure does not bear on the compensatory account discussed here (which concerns supraglottal gestures) and laryngeal gestures can be implemented almost entirely independently from oral gestures, and (2) the voicing effect has been shown to exist even in whispered speech, where vocal fold vibration is entirely absent (Sharf 1964).<sup>7</sup>

The durations in milliseconds of the following intervals were extracted with a series of custom Praat scripts from the annotated acoustic landmarks: word duration, vowel duration (V1 onset to V1 offset), consonant closure duration (V1 offset to C2 release), and release to release duration (C1 release to C2 release). Sentence duration was measured in seconds. Figure 1 shows an example of the segmentation of /pata/ (a) and /pada/ (b) from an Italian speaker. Syllable rate (syllables per second) was used as a proxy to speech rate (Plug & Smith 2018), and was calculated as the number of syllables divided by the duration of the sentence in seconds (8 syllables in Italian, 6 in Polish). All further data processing and visualisation was done in R v3.5.2 (R Core Team 2018; Wickham 2017).

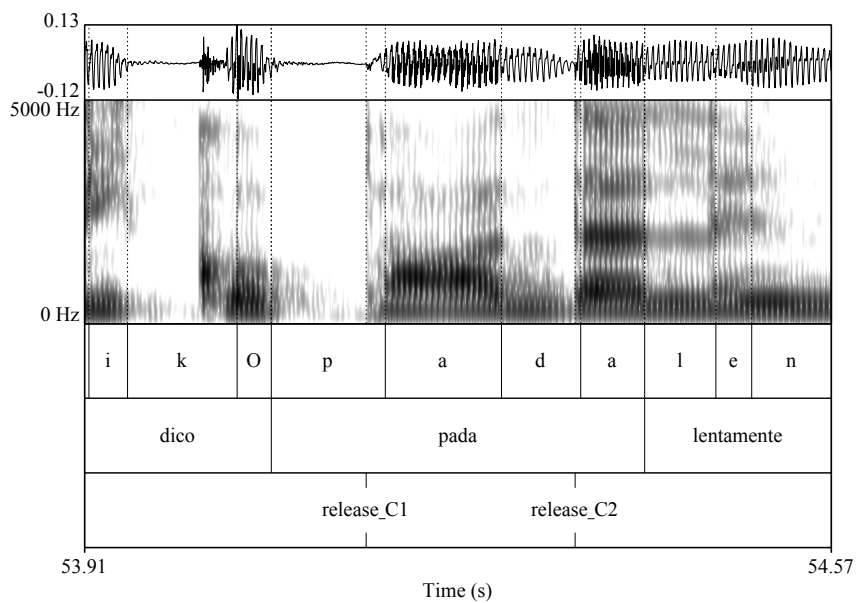
## 2.6 Statistical analysis

Given the data-driven nature of the study, all statistical analyses reported here are to be considered exploratory (hypothesis-generating) rather than confirmatory (hypothesis-driven, Kerr 1998; Gelman & Loken 2013; Roettger 2018). The durational measurements were analysed with linear mixed-effects models using lme4 v1.1-19 in R (Bates et al. 2015), and model estimates were extracted with the effects package v4.1-0 (Fox 2003). All factors were coded with treatment contrasts and the following reference levels: voiceless (vs. voiced), /a/ (vs. /o/, /u/), coronal (vs. velar), Italian (vs. Polish). Speech rate has been centred when included in the models to make the intercept estimates more interpretable. The models were fit-

<sup>7</sup> A reviewer makes interesting phonological remarks. The presence of lenition and voicing of voiceless stops in some varieties of Italian and its absence in Polish could be related to differences in laryngeal phonology and prosodic structure between these languages, namely the absence of a feature [voice] in Italian and the absence of true trochees in Polish. This hypothesis is compatible with work by Schwartz & Arndt (2018) and Schwartz (2016), to which the reader is referred.



(a) /pata/



(b) /pada/

**Figure 1:** Segmentation example of the words *pata* and *pada* uttered by the Italian speaker IT09 (the times on the x-axis refer to the times in the concatenated audio file).

ted by Restricted Maximum Likelihood estimation (REML). The estimates in the results section refer to these reference levels unless interactions are discussed. *P*-values for the individual terms were obtained with `lmerTest` v3.0-1, which uses the Satterthwaite's approximation to degrees of freedom (Kuznetsova et al. 2017; Luke 2017). A result is considered significant if the *p*-value is below the alpha level ( $\alpha = 0.05$ ). The choice of not using likelihood ratio tests for statistical inference is based on Luke (2017) who argues that this approach can lead to inflated Type I error rates. In any case, Luke (2017: 1501) also warns that 'results [from mixed-effects models] should be interpreted with caution, regardless of the method adopted for obtaining *p*-values'. Inspection of residual plots and QQ plots of the models described below indicated absence of patterns in the residuals.

Bayes factors were used to test whether word and release to release duration are not affected by C2 voicing (i.e., the effect of C2 voicing on duration is 0).<sup>8</sup> For each set of null/alternative hypotheses, a full model (with the predictor of interest) and a null model (excluding it) were fitted separately using the Maximum Likelihood estimation (ML, Bates et al. 2015: 34). The Bayes Information Criterion (BIC) approximation was then used to obtain Bayes factors (Raftery 1995; 1999; Wagenmakers 2007; Jarosz & Wiley 2014). The approximation is calculated according to the equation in 1 (Wagenmakers 2007: 796).

$$(1) \quad BF_{01} \approx \exp(\Delta BIC_{10}/2)$$

where  $\Delta BIC_{10} = BIC_1 - BIC_0$ ,  $BIC_1$  is the BIC of the full model, and  $BIC_0$  is the BIC of the null model. Values of  $BF_{01} > 1$  indicate a preference of  $H_0$  over  $H_1$ . The interpretation of the Bayes factors follows the recommendations in Raftery (1995: p. 139): 1–3 = weak evidence, 3–20 = positive evidence, 20–150 = strong evidence,  $> 150$  = very strong evidence.

The extracted measurements were filtered before statistical analysis. Measures of vowel duration, closure duration, word duration, and release to release duration that are 3 standard deviations lower or higher than the

<sup>8</sup> The choice of Bayes factors over other information criteria, like AIC, is a practical one. First, Bayes factors can be used to identify the relative strength of the evidence for each hypothesis. The higher the Bayes factor of  $H_{01}$ , the stronger the evidence for  $H_0$  according to the data. Second, a Bayes factor near 1 indicates that the data is compatible with both hypotheses (even when AIC indicates a preference of one over the other), in which case it is not possible to choose among them. Note that the AICs of the word duration and release to release duration models reported below are lower when C2 voicing is not included as a predictor than when it is included, although the difference in AIC between the null and full models is very small (below 2).

respective means were excluded from the final dataset (this procedure generally corresponds to a loss of around 2.5% of the data). One sentence (sentence 48 of IT07, *Dico pada lentamente*) included a speech error and has been excluded. After excluding missing measurements, these operations yield a total of 920 tokens of vowel and closure durations, 1176 tokens of word duration, and 848 tokens of release to release duration.

## 2.7 Open Science statement

Following recommendations for Open Science in Crüwell et al. (2018) and Berez-Kroeker et al. (2018), the data and code used to produce the analyses discussed in this paper are available on the Open Science Framework at <http://bit.ly/2FitZe1>.

## 3 Results

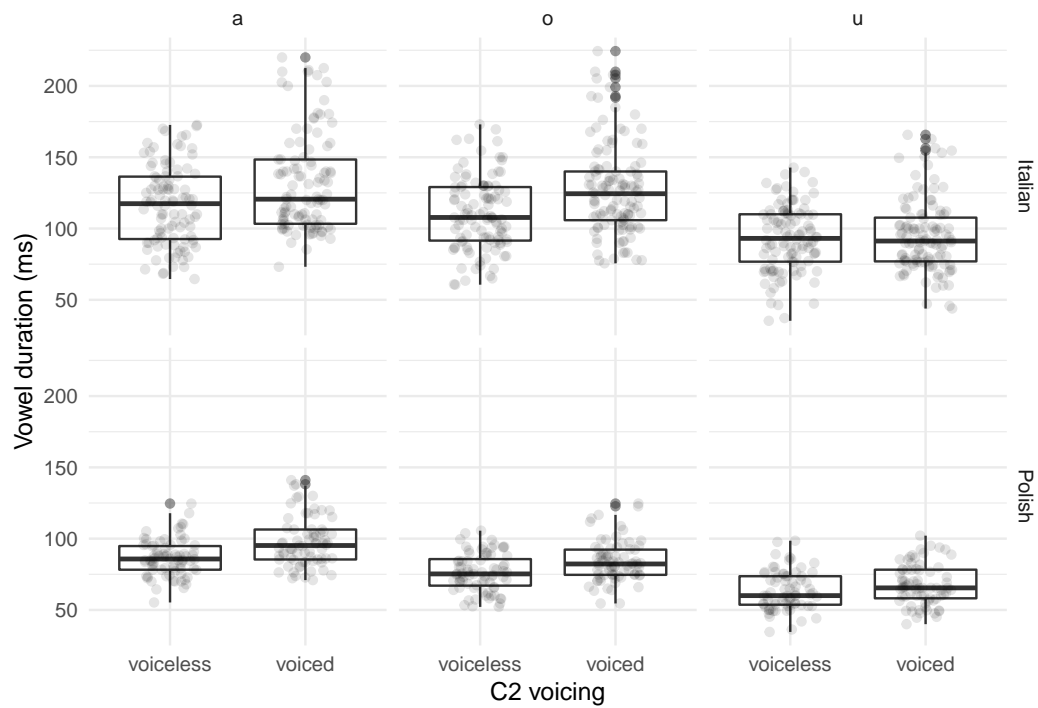
The following sections report the results of the study in relation to the durations of vowels, consonant closure, word, and the release to release interval. When discussing the output of statistical modelling, only the relevant predictors and interactions will be presented. The full output of statistical models (including confidence intervals and *p*-values) are given in Appendix A.

### 3.1 Vowel duration

Figure 2 shows boxplots and raw data of vowel duration for the three vowels /a, o, u/ when followed by voiceless or voiced stops in Italian and Polish. Vowels tend to be longer when followed by a voiced stop in both languages. The effect appears to be greater in Italian than in Polish, especially for the vowels /a/ and /o/. There is no evident effect of C2 voicing in /u/ in Italian, but the effect is discernible in Polish /u/. In Italian, vowels have a mean duration of 106.16 ms (SD = 27.08) before voiceless stops, and a mean duration of 117.66 ms (SD = 34.63) before voiced stops. Polish vowels are on average 75.57 ms long (SD = 16.16) when followed by a voiceless stop, and 83.11 ms long (SD = 19.37) if a voiced stop follows. The difference in vowel duration based on the raw means is 11.5 ms in Italian and 7.54 ms in Polish.

A linear mixed-effects model with vowel duration as the outcome variable was fitted with the following predictors: fixed effects for C2 voicing (voiceless, voiced), C2 place of articulation (coronal, velar), vowel (a, o,





**Figure 2:** Raw data and boxplots of the duration in milliseconds of vowels in Italian (top row) and Polish (bottom row), for the vowels /a, o, u/ when followed by a voiceless or voiced stop.

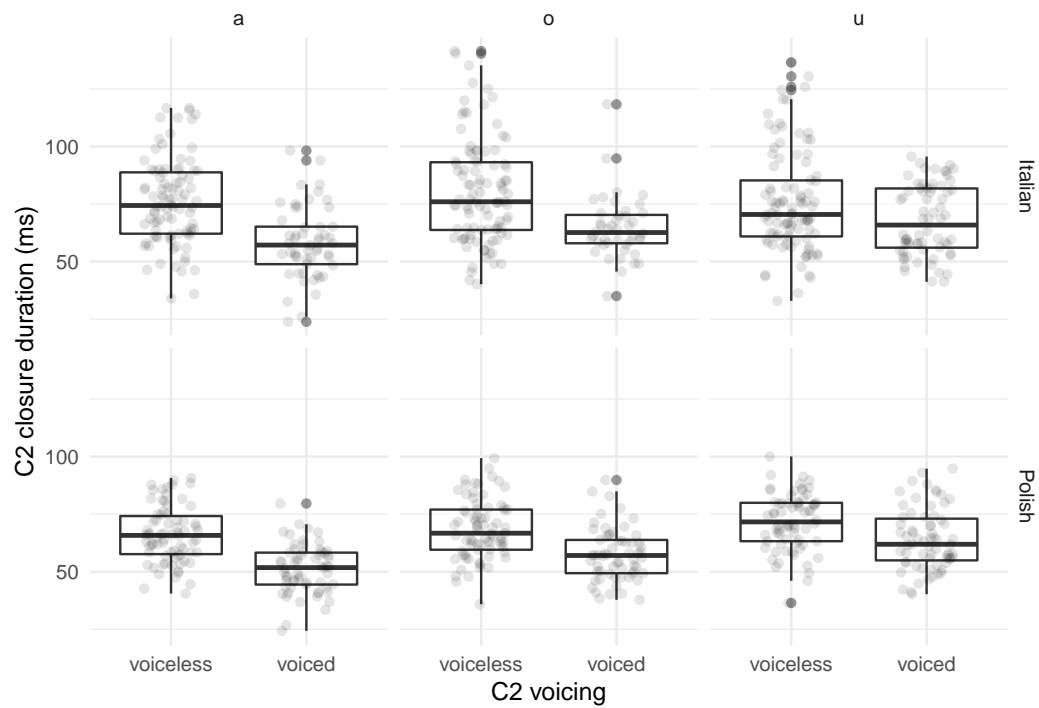
u), language (Italian, Polish), and speech rate (as syllables per second, centred); by-speaker and by-word random intercepts with by-speaker random slopes for C2 voicing. All possible interactions between C2 voicing, vowel, and language were included. The following terms are significant according to *t*-tests with Satterthwaite's approximation to degrees of freedom: C2 voicing, C2 place, vowel, language, and speech rate. Only the interaction between C2 voicing and vowel is significant. Vowels are 16.28 ms longer (SE = 4.42) when followed by a voiced stop (C2 voicing), and 8 ms shorter (SE = 1.63) when followed by a velar stop. The effect of C2 voicing is smaller with /u/ (around 3 ms,  $\hat{\beta} = -13.1$  ms, SE = 5.56). Polish has on average shorter vowels than Italian ( $\hat{\beta} = -24.05$  ms, SE = 7.83), and the effect of voicing is estimated to be about 10.55 ms (although note that the interaction between language and C2 voicing is not significant). Speech rate has a negative effect on vowel duration, such that faster rates correlate with shorter vowel durations ( $\hat{\beta} = -16.23$  ms, SE = 1.26).

### 3.2 Consonant closure duration

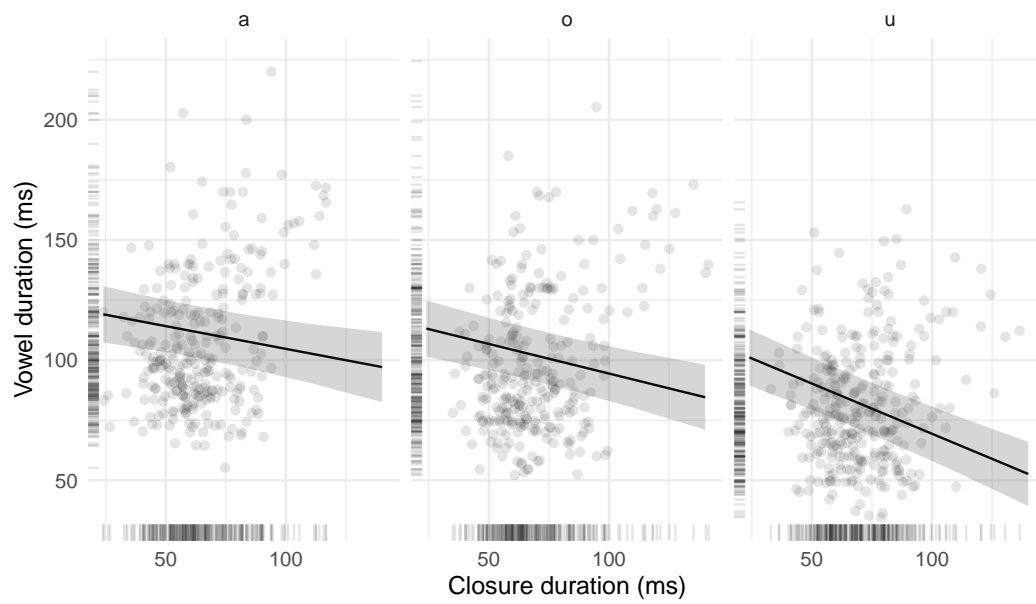
Figure 3 illustrates stop closure durations with boxplots and individual raw data points. A pattern opposite to that with vowel duration can be noticed: closure duration is shorter for voiced than for voiceless stops. The closure of voiced stops in Italian is 106.16 ms long (SD = 27.08), while the voiceless stops have a mean closure duration of 117.66 ms (SD = 34.63). In Polish, the closure duration is 75.57 ms (SD = 16.16) in voiced stops and 83.11 ms (SD = 19.37) in voiceless stops. The difference in closure duration based on the raw means is 13.33 ms in Italian and 10.87 ms in Polish. The same model specification as with vowel duration has been fitted with consonant closure duration as the outcome variable. C2 voicing, C2 place, and speech rate are significant. Stop closure is 17.5 ms shorter (SE = 4) if the stop is voiced and 3.5 ms longer (SE = 1.5) if velar. Finally, faster speech rates correlate with shorter closure durations ( $\hat{\beta} = -8.5$  ms, SE = 1 ms).

### 3.3 Vowel and closure duration

A model addressing the relationship between vowel and stop closure duration was fitted with the following terms and interactions: vowel duration as the outcome variable; as fixed effects, closure duration, vowel, speech rate (centred); all logical interactions between closure duration, vowel, and speech rate; by-speaker and by-word random intercepts. Closure duration



**Figure 3:** Raw data and boxplots of closure duration in milliseconds of voiceless and voiced stops in Italian (top row) and Polish (bottom row) when preceded by the vowels /a, o, u/.



**Figure 4:** Raw data, estimated regression lines, and 95 per cent confidence intervals of the effect of closure duration on vowel duration for the vowels /a, o, u/ (from a mixed-effects model fitted to data pooled from Italian and Polish, see text for details).

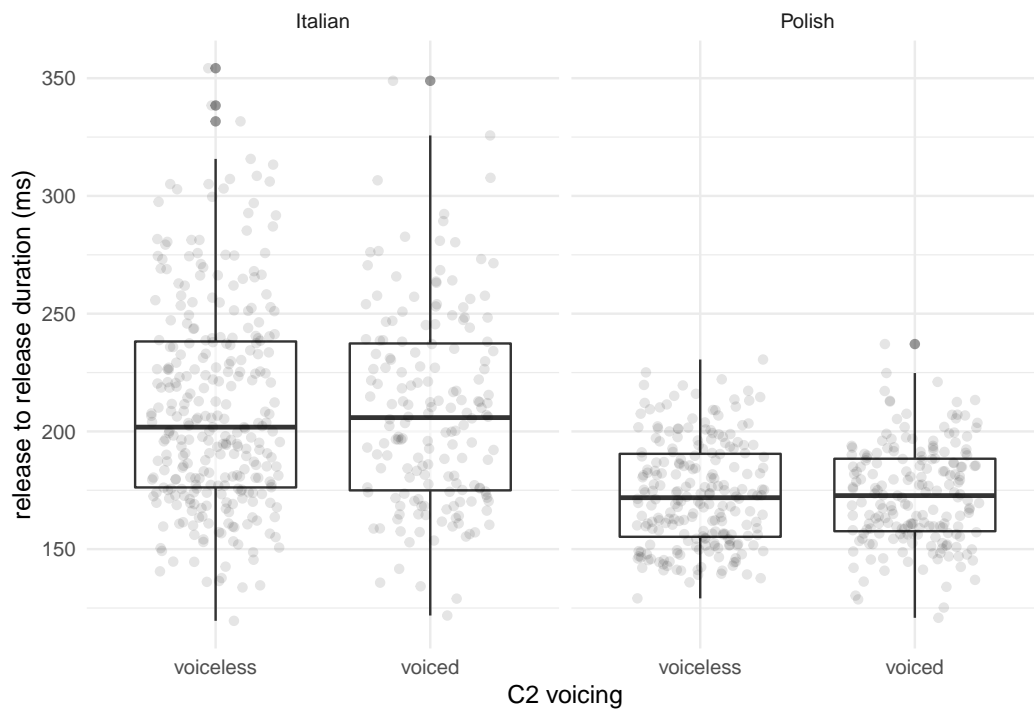
has a significant effect on vowel duration ( $\hat{\beta} = -0.19$  ms, SE = 0.06 ms). The effect with /u/ is greater than with /a/ and /o/ ( $\hat{\beta} = -0.23$  ms, SE = 0.08 ms). In general, closure duration is inversely proportional to vowel duration. However, such a correlation is quite weak, as shown by the small estimates. A 1 ms increase in closure duration corresponds to a 0.2–0.45 ms decrease in vowel duration. These estimates can be interpreted in terms of percentages of compensation, which range between 20 and 45%. Note, moreover, that the negative correlation found here could be a consequence of annotation bias, since the vowel and closure share a boundary. Faster speech rates elicit a bigger effect than lower speech rates, as indicated by the significant interaction between closure duration and speech rate ( $\hat{\beta} = -0.2$  ms, SE = 0.06 ms). The effect of the interaction is reduced when the vowel is /u/ ( $\hat{\beta} = 0.17$  ms, SE = 0.08 ms). Figure 4 shows for each vowel /a, o, u/ the individual data points and the regression lines with 95% confidence intervals extracted from the mixed-effects model.

### 3.4 Word duration

Words with a voiceless C2 are on average 393.72 ms long (SD = 79.05) in Italian and 387.72 ms long (SD = 73.45) in Polish. Words with a voiced stop have a mean duration of 357.07 ms (SD = 39.14) in Italian and 361.87 ms (SD = 38.51) in Polish. The following full and null models were fitted to test the effect of C2 voicing on word duration. The full model is made up of the following fixed effects: C2 voicing, C2 place, vowel, language, and speech rate. The model also includes by-speaker and by-word random intercepts, and a by-speaker random slope for C2 voicing. The null model is the same as the full model with the exclusion of the fixed effect of C2 voicing. The Bayes factor of the null against the full model is 19. Thus, the null model (in which there is no effect of C2 voicing,  $\beta = 0$ ) is 19 times more likely under the observed data than the full model. This indicates that there is positive evidence for a null effect of C2 voicing on word duration.

### 3.5 Release to release interval duration

In Figure 5, boxplots and raw data points show the duration of the release to release interval in words with a voiceless vs. a voiced C2 stop, in Italian and Polish. It can be seen that the distributions, medians, and quartiles of the durations in the voiceless and voiced condition do not differ much in either language. In Italian, the mean duration of the release to release interval is



**Figure 5:** Raw data and boxplots of the duration in milliseconds of the release to release interval in Italian (left) and Polish (right) when C2 is voiceless or voiced.



209.88 ms (SD = 43.84) if C2 is voiceless, and 208.6 ms (SD = 41.34) if voiced. In Polish, the mean durations are respectively 173.13 (SD = 22.44) and 172.67 (SD = 20.47) ms. The specifications of the null and full models for the release to release duration are the same as for word duration. The Bayes factor of the null model against the full model is 21, which means that the null model (without C2 voicing) is 21 times more likely than the model with C2 voicing as a predictor. The Bayes factor suggests there is strong evidence that duration of the release to release interval is not affected by C2 voicing.

## 4 Discussion

A study of articulatory and acoustic aspects of the effect of consonant voicing on vowel duration in Italian and Polish has been carried out to look for a possible source of such an effect in speech production. Only the results from the acoustic part of the study bear on the main argument of this paper. The following sections discuss, in turn, the results regarding the effect of voicing on vowel duration in Italian and Polish and how the finding that the duration of the interval between the two consecutive consonant releases in CVCV words is compatible with a compensatory temporal adjustment account of the voicing effect. The section concludes by discussing the limitations and open issues of this study.

### 4.1 Voicing effect in Italian and Polish

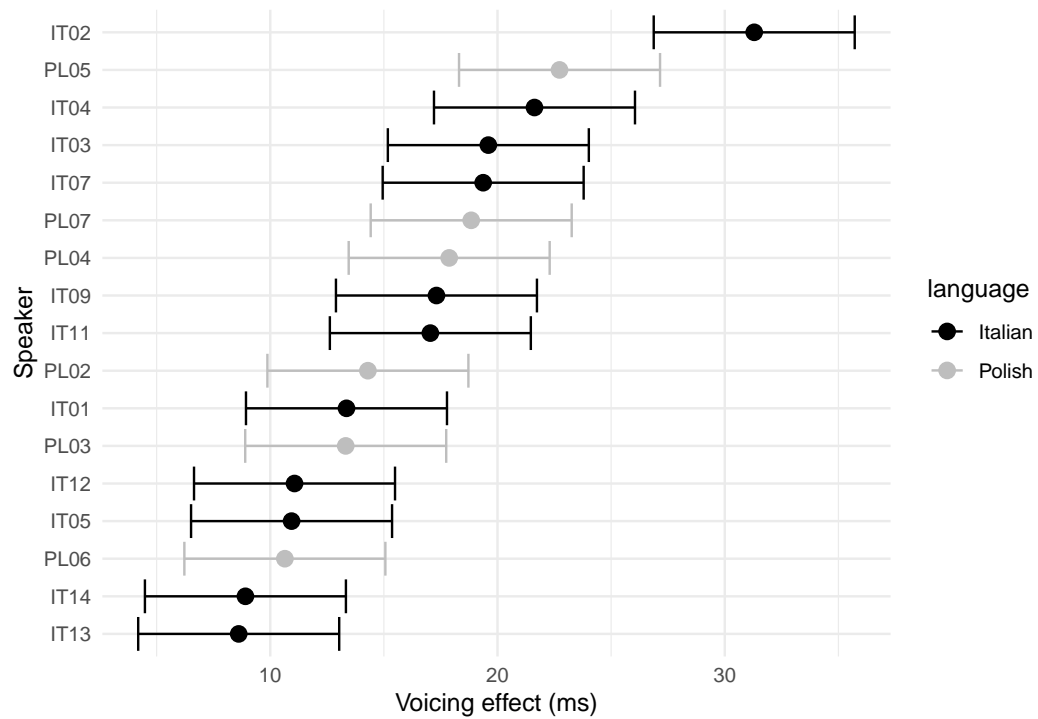
The results of vowel duration and C2 voicing indicate that vowels are longer when followed by voiced than when followed by voiceless stops both in Italian and Polish. The estimated effect is around 16 ms when C2 is voiced for Italian. This value is not too far from the estimates of previous works on this language (Magno Caldognetto et al. 1979; Farnetani & Kori 1986; Esposito 2002), the range of which is between 22 and 24 ms. The higher estimates of these studies compared to the one here could be related to differences in experimental design, or Type M (magnitude) errors due to low statistical power (see Kirby & Sonderegger 2018). The estimate of the effect of voicing on C2 closure duration is around -18 ms. Crucially, the effect of voicing on vowel and closure duration have very similar magnitudes and opposite signs. These results suggest a compensatory mechanism between vowel and closure duration.

Furthermore, the effect of voicing on the duration of Italian /u/ is smaller than with /a/ and /o/ (about 3 vs. 16 ms respectively), a fact already observed by Ferrero et al. 1978. While it is not clear why the duration of this particular vowel should not be affected by C2 voicing, the data reported here indicate that the magnitude of the difference in closure duration when the preceding vowel is /u/ is smaller than with /a/ and /o/ (about 7 vs. 17 ms respectively). If vowel duration compensates for closure duration, then a smaller difference in closure duration should correspond to a small difference in vowel duration, as the estimates seem to suggest.

The interpretation of the Polish results is less straightforward. Previous studies found either no voicing effect or a small effect in Polish (3.5–4.5 ms). In particular, Malisz & Klessa (2008) say that the effect seems to be very idiosyncratic in the 40 speakers of their analysis. The estimated effect found in the 6 Polish speakers of the present study is about 10.5 ms, and the difference based on the means of the raw vowel durations is 7.5 ms. Recall, however, that the interaction between language and C2 voicing (which gives the estimate of 10.54) is not significant (see the full model summary in Table 2). It is likely, though, that the non-significance might be related to low power. Indeed, the raw mean difference of 7.5 ms in Polish—although still higher than what found in previous studies—might be more informative.

More specifically, when one compares the raw mean duration differences of vowels with the raw mean duration differences of consonant closures, a pattern can be seen. The mean differences of Italian vowels and closures (11.5 and 13.33, respectively) are bigger than those of Polish (7.54 and 10.87), even if by just a small amount. It is plausible that the smaller effect of C2 voicing on preceding vowel duration in Polish is related to the smaller effect on closure duration, if we assume a temporal mechanism of compensation between the closure and the vowel. These patterns will need to be confirmed with a more balanced sample of Italian and Polish speakers.

On the other hand, while the estimated differences in vowel durations can be interpreted in reference to Italian and Polish as two independent linguistic objects, the patterns observed in the individual speakers does not indicate a systematic relation between magnitude of the effect and language. Figure 6 shows the random coefficients of the effect of C2 voicing on vowel duration for the individual speakers, extracted from the mixed-effects model presented in Section 3.1. Black indicates Italian speakers, while grey is for Polish speakers. As can be seen, speakers of both languages are scattered along the values of the voicing effect. These results are in agreement with the idiosyncrasy of the voicing effect of Polish found in Malisz & Klessa



**Figure 6:** By-speaker random coefficients and error bars for the effect of C2 voicing on vowel duration, extracted from a mixed-effect model (Section 3.1).

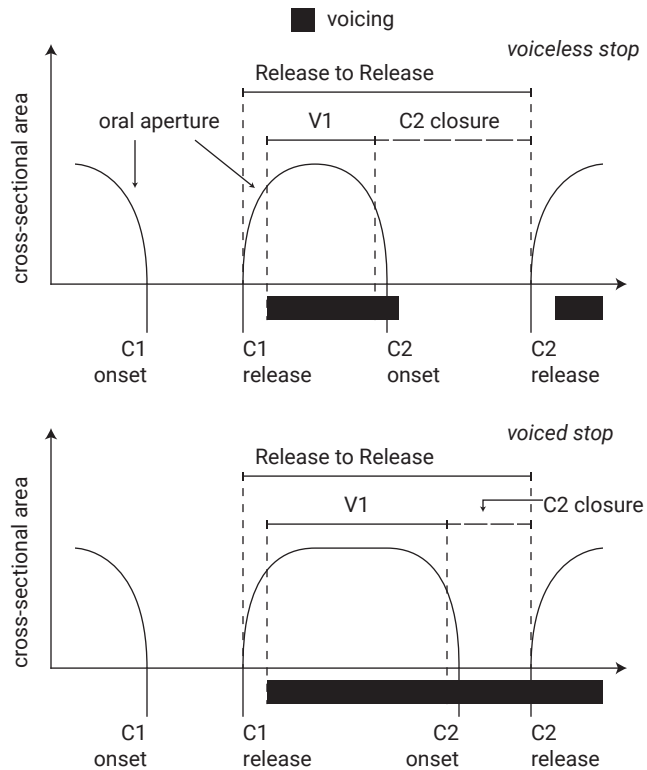
(2008). While large-scale studies could reveal clear language-level patterns, the data discussed here point to a scenario in which the speaker's individual behaviour is substantial. Future studies could thus look into the respective role of individual-level and community-level factors and how these contribute to the magnitude of the durational differences across speakers and languages.

## 4.2 *Compensatory temporal adjustment*

Vowels followed by voiced stops are long, while vowels followed by voiceless stops are short. The closure duration of voiced stops is short compared to that of voiceless stops. There seems to be an inverse relation between vowel duration and closure duration, by which a long vowel entails a short closure (and vice versa), and a short vowel entails a long closure (and vice versa).

The data and statistical analyses of this study suggest that the duration of the interval between the releases of two consecutive consonants in CVCV words (the release to release interval) is not affected by the phonological voicing of the second consonant (C2) in Italian and Polish. In accordance with a compensatory temporal adjustment account (Slis & Cohen 1969b; Lehiste 1970b), the difference in vowel duration and closure durations before voiceless vs. voiced stops can be seen as the outcome of differences in timing of the vowel offset/closure onset (more neutrally, the VC boundary). In other words, the timing of the VC boundary within the temporally stable release to release interval determines the duration of both the vowel and the stop closure. An earlier VC boundary relative to the onset of the preceding vowel results in a shorter vowel and a longer stop closure. On the other hand, a later VC boundary produces a longer vowel and a shorter closure. Figure 7 illustrates this compensatory mechanism. Note that the term 'temporal stability' (and 'temporally stable') as used here means that the underlying statistical distribution of the interval duration is stable *across contexts of C2 voicing*. No specific statement is implied about the variance of the duration around the mean, across or within phonological contexts.

The invariance of the release to release interval allows us to refine the logistics of the compensatory account by narrowing the scope of the temporal adjustment action. A limitation of this account, as proposed by Slis & Cohen (1969b) and Lehiste (1970b), is the lack of a precise identification of the word-internal mechanics of compensation. As already discussed in Section 1, it is not clear why the adjustment should target the preced-



**Figure 7:** A schematic representation of the oral cavity cross-sectional area, as inferred from acoustics. Design based on [Esposito \(2002\)](#). The top panel shows a CVC sequence with a voiceless C2, the bottom panel with a voiced C2. Oral cavity aperture (on the y-axis, as the inverse of oral constriction) through time (on the x-axis) is represented by the black line.

Lower values represent a more constricted oral tract (a contoid configuration), while higher values indicate a more open oral tract (a vocoid configuration). The black bars below the time axis represent voicing (vocal fold vibration). Various landmarks and intervals are indicated in the schematic.

ing stressed vowel, rather than the following unstressed vowel or any other segment in the word. Since the release to release interval includes just the vocoid gesture between the release of C1 and the VC boundary, and the consonant closure, it follows that differences in the timing of the VC boundary must be reflected in differences in both vowel and closure durations.

Under an account of temporal compensation, the voicing effect can be interpreted as a by-product of gestural phasing and mechanisms operating on the timing of the VC boundary. The temporal stability of the release to release interval across voicing contexts allows us to refine the compensatory mechanism by providing a temporal anchor. On the other hand, it is important to note that the release to release interval should not necessarily have a special status in such a compensatory account, but rather can be used as a proxy to the understanding of a full gestural mechanism of compensation. Indeed, the temporal stability of this interval should be derivable from a theory of gestural phasing, rather than one that simply states that the interval is stable across voicing contexts.

The non-exclusivity of the release to release interval is also shown by the fact that excluding the VOT from it still indicates that C2 voicing is not affecting the interval duration. The duration of the vowel onset to release interval (the release to release minus VOT) is stable across voicing contexts (Bayes factor = 9). However, the duration of release to release interval has relatively more cohesion than that of the vowel onset to release interval, as indicated by two measures of relative dispersion (the coefficient of variation CV and the coefficient of quartile variation CQV, see [Bonett 2006](#)).<sup>9</sup> On the other hand, the duration of the interval between the vowel onset of V1 to the vowel onset of V2 does change depending on C2 voicing (the interval it's around 20 ms longer if C2 is voiceless). This fact is simply a consequence of including the VOT of C2 in the measure. Voiceless stops have longer VOT values, which increases the duration of the interval. The difficulty in identifying a clear-cut time point corresponding to vowel onset could explain the relative higher dispersion of the vowel onset to release interval duration. For these reasons, the release to release interval is probably a better measure of temporal stability than the vowel onset to release, given its inherent higher cohesion.

It is possible that the temporal stability of the release to release interval is not an antecedent, but rather a consequence of manipulating vowel and closure durations. If this were the case, the differential duration of vowels

<sup>9</sup> The CV of the release to release duration is 0.203, while that of the vowel onset to release duration is 0.232. The CQV is 0.127 for the release to release and 0.136 for the vowel onset to release. Lower values mean less dispersion/more cohesion.



and closures would not be the result of a compensatory mechanism. The present data cannot disambiguate between these two scenarios, and future studies should look into investigating independent reasons for the release to release interval stability across voicing contexts. The account of gestural phasing proposed by Tilsen (2013; 2016) is promising, in that the temporal stability of the release to release interval would directly follow from the relative phasing of the vowels in CVCV words (see Figure 6 in Tilsen 2013). Articulatory work on the gestural coordination of sequences besides the traditional syllable might reveal a principled organisation that results in the temporal patterns observed in this study and in other durational phenomena.

However, even if independent reasons for the interval stability can be identified, other mechanisms, unrelated to compensatory effects, would still be required to explain the differential timing of the VC boundary within that interval. Accounts compatible with other aspects of production and perception would not be ultimately ruled out, as thoroughly discussed in Beguš (2017). For example, the laryngeal adjustment hypothesis (Halle et al. 1967) states that adjustments of the glottis for obstruent voicing require more time to be implemented, so that stop closure onset (VC boundary) for a voiced stops will be achieved later than that of a voiceless stop, relative to the onset of the preceding vowel. Tongue root advancement (Rothenberg 1967; Westbury 1983; Ohala 2011) could also play a role in modulating the time required before closure can be implemented. Another account (Chen 1970) makes direct reference to velocity of the closing gesture, which is faster in voiceless than in voiced stops (Van Summers 1987; de Jong 1991), so that the VC boundary within the release to release would be timed earlier in the former than in the latter case. Moreover, perceptual explanations of the voicing effect have been proposed in Javkin (1976) and Kluender et al. (1988), and these perceptual factors might play a role in the enhancement of the effect (see Kingston & Diehl 1994, Port & Dalby 1982, and Luce & Charles-Luce 1985, see Fowler 1992 for a critique to Kluender et al. 1988). Finally, whether the timing of the VC boundary depends on modulations of the vocalic or consonantal gesture, or both, is another aspect that should be investigated further (see de Jong 1991 for an example).

A comment is also due in relation to possible coexisting effects on vowel duration. Beguš (2017) finds that, even when C2 closure duration is controlled for, C2 phonation (ejective, voiceless, voiced) in Georgian is still a significant predictor. The author argues for a separate laryngeal features effect, which operates in addition to a closure duration effect. In the present study, C2 voicing (voiceless, voiced) and its interactions are not significant

when included in the model discussed in section Section 3.3, which has vowel duration as outcome and C2 closure duration as one of the predictors.<sup>10</sup> However, even when multicollinearity between predictors is minimal, presence or lack of statistical significance of multiple terms cannot unequivocally inform us on the actual contribution of those terms, since it is possible that unknown relations between terms mask underlying mechanisms (for a discussion see [McElreath 2015](#)). The diachronic development of context-driven statistical sub-distributions can override the original causal link ([Sóskuthy 2013](#)). Under this scenario, it is not possible to discern which of the competing predictors is diachronically responsible for the relation, and either or both the compensatory mechanism and the laryngeal features could have had a role in generating the synchronic patterns (this kind of reasoning is compatible for example with exemplar theories of speech perception and production, see among others [Johnson 1997](#); [Sóskuthy et al. 2018](#); [Ambridge 2018](#); [Todd et al. 2018](#)).

Since diverging results have been obtained in relation to the significance of C2 phonation in addition to C2 closure durations, these aspects need to be further investigated in future studies, although to ascertain whether they are artefacts of statistical procedures or if they reflect an underlying state of affairs might still prove difficult. To conclude, lack of significance of a separate laryngeal features effect in this study cannot be taken as evidence for its absence in the present data, and a compensatory mechanism could coexist with mechanisms directly related to laryngeal features, which would in turn explain the differential timing of the VC boundary.

### 4.3 *Limitations and future work*

The generalisations put forward in this paper strictly apply to disyllabic words with a stressed vowel in the first syllable, flanked by single stops. First, it is possible that the pattern found in this context does not occur in sequences including an unstressed vowel. For example, it is known that the difference in closure duration between voiceless and voiced stops is not stable when the stops precede a stressed vowel, although vowels preceding pre-stress stops have slightly different durations ([Davis & Van Summers 1989](#)). According to the mechanism proposed here, the absence of differences in closure duration should correspond to the absence of differences in vowel duration. Second, it is known that the magnitude of the effect of voicing is modulated by other prosodic characteristics, like the number of syllables

<sup>10</sup> Multicollinearity is not an issue here, since the VIFs are all below 3 ([Zuur et al. 2010](#)).

in the word, presence/absence of focus, and position within the sentence (Sharf 1962; Klatt 1973; Laeufer 1992; de Jong 2004). Third, the constraints on experimental material enforced by the use of ultrasound tongue imaging have been previously mentioned in Section 2.3. Given these constraints, temporal information from other vowels (like front vowels), places and manners of articulation is a desideratum. Data from different contexts and different languages is thus needed to assess the generality of the claims put forward in this paper.

Another issue is the interaction of the temporal compensation and speech rate. The magnitude of compensation between vowel and closure duration found in de Jong (1991) and here is somewhat small (between 12% and 40%). Ideally, given the temporal stability of the release to release interval relative to C2 voicing, the compensation rates should approximate 100%. However, it is possible that the correlation between vowel and closure duration is modulated in complex ways by the individual effects of speech rate on the vowel and the closure. For example, Ko (2018) finds that the vowel/closure ratio differs depending on speaking rate and that there is an interaction between the voicing of the consonant and speaking rate. When the consonant is voiceless, the vowel/closure ratio is smaller when speaking rate is slow, while slow speaking rate induces larger vowel/closure values when the consonant is voiced. Experimental work is required which addresses the differential effect of speaking rate on vowel and consonant closures, and how these interact with a possible compensatory mechanism.

Some concern could be raised in relation to possible influences of English on the native productions of participants recorded in the English-speaking context of the University of Manchester Laboratory. However, as reported in Section 2.4, conversations during the session prior to the experiment and instructions were in the participant's native language. Antoniou et al. (2010) show that, in a situational language context study of Greek-English bilinguals, being exposed to the native language during the experiment elicited Greek native-like phonetic values even when the dominant language at the time of recording was English (the bilingual speakers acquired English as a second language, being Greek their first). A small effect of L2 could persist in proficient L2 speakers, as found by Schwartz et al. (2015). The five Polish speakers with a highly proficient level of English investigated in that study showed a 10 ms increase in VOT values compared to the quasi-monolingual base level. While previous studies focussed on VOT, future work should directly test the influence of English on the magnitude of the voicing effect of one's native language.

The compensatory temporal adjustment account presented here extends to other durational effects discussed in the literature. In particular, the account bears predictions on the direction of the durational difference led by phonation types different from voicing, like aspiration and ejection. For example, the mix of results with regard to the effect of aspiration (Durvasula & Luo 2012) suggests that the conditions for a temporal adjustment might differ across the contexts and languages studied. In light of the results in Beguš (2017), future studies will also have to investigate the durational invariance of speech intervals in relation to a variety of phonation contrasts.

## 5 Conclusions

The results of this exploratory study of the effect of voicing on vowel duration are congruent with a compensatory temporal adjustment account of such effect. Acoustic data from seventeen speakers of Italian and Polish show that the temporal distance between two consecutive stop releases is not affected by the voicing of the second stop in CVCV words. The temporal invariance of the release to release interval, together with a difference in timing of the VC boundary, can cause vowels to be shorter when followed by voiceless stops (which have a long closure) and longer when followed by voiced stops (the closure of which is short).

As discussed in Section 4.2, the temporal patterns reported here do not univocally exclude other possible sources for the duration differential. Multiple mechanisms (both articulatory and perceptual) could conspire together to produce the observed patterns. Such a pluralist view has already been proposed for the voicing effect (for example, Beguš 2017 and Sanker 2018), and for other related phenomena, like vowel duration in incomplete neutralisation (Winter & Roettger 2011). For a review of explanatory pluralism in the cognitive sciences, see Dale et al. 2009 and references therein. Indeed, a hybrid account, which takes into consideration and synthesises aspects of multiple proposed accounts, is probably warranted, given the diversity of compatible results obtained so far. Future work will need to investigate further aspects of the patterns found in this study, with a particular focus on the effects of different segmental and prosodic structures and different laryngeal contrasts on the release to release interval, and in relation to other attributes of consonant effects on vowel duration.

## 777 **Acknowledgements**

778 I am grateful to Ricardo Bermúdez-Otero and Patrycja Strycharczuk for  
779 their immense support and patience in providing feedback on this project.  
780 Thanks to the editor of Glossa and two anonymous reviewers for their feed-  
781 back, which significantly improved the manuscript. I also want to thank  
782 the audience at the 16th Laboratory Phonology conference (LabPhon16)  
783 for their input, and Kenneth de Jong for comments on an early draft of  
784 this paper. Thanks also go to my colleagues at the Phonetics Laboratory  
785 of the University of Manchester, who provided help in different ways. Any  
786 remaining errors are my own.

## 787 **Funding information**

788 This project has been funded by the School of Arts, Languages, and Cultures  
789 Graduate School at the University of Manchester.

## 790 **Competing interests**

791 The author has no competing interests to declare.

## 792 **A Output of statistical models**

793 See Table 2, Table 3, and Table 4.

## 794 **B Socio-linguistic information of participants**

795 See Table 5.

## 796 **C Target words**

797 See Table 6.

**Table 2:** Summary of the linear mixed-effects model fitted to vowel duration (see Section 3.1).

Predictor	Estimate	SE	CI low	CI up	df	t-value	p-value	< $\alpha$
Intercept	118.06	4.94	108.38	127.74	23.89	23.91	0.00	*
Voicing = voiced	16.28	4.42	7.62	24.95	15.39	3.68	0.00	*
Vowel = /o/	-7.50	3.93	-15.21	0.21	10.31	-1.91	0.08	
Vowel = /u/	-25.71	3.94	-33.44	-17.98	10.43	-6.52	0.00	*
Lang = Polish	-24.05	7.83	-39.40	-8.70	22.38	-3.07	0.01	*
Place = velar	-7.95	1.63	-11.15	-4.75	10.99	-4.87	0.00	*
Speech rate	-16.23	1.26	-18.70	-13.77	854.64	-12.92	0.00	*
Voiced $\times$ /o/	2.09	5.54	-8.77	12.96	10.18	0.38	0.71	
Voiced $\times$ /u/	-13.09	5.56	-23.99	-2.20	10.30	-2.36	0.04	*
Voiced $\times$ Polish	-5.73	6.61	-18.69	7.23	18.00	-0.87	0.40	
/o/ $\times$ Polish	-2.50	5.66	-13.60	8.60	11.09	-0.44	0.67	
/u/ $\times$ Polish	1.12	5.68	-10.01	12.26	11.23	0.20	0.85	
Voiced $\times$ /o/ $\times$ Polish	-6.16	8.00	-21.85	9.53	11.06	-0.77	0.46	
Voiced $\times$ /u/ $\times$ Polish	6.40	8.03	-9.34	22.13	11.19	0.80	0.44	

**Table 3:** Summary of the linear mixed-effects model fitted to closure duration (see Section 3.2).

Predictor	Estimate	SE	CI low	CI up	df	t-value	p-value	< $\alpha$
Intercept	73.25	4.28	64.86	81.63	22.38	17.11	0.00	*
Voicing = voiced	-17.70	4.06	-25.66	-9.74	18.63	-4.36	0.00	*
Vowel = /o/	3.75	3.26	-2.64	10.13	9.43	1.15	0.28	
Vowel = /u/	-1.91	3.27	-8.32	4.49	9.56	-0.58	0.57	
Lang = Polish	-7.03	6.82	-20.40	6.34	20.82	-1.03	0.31	
Place = velar	3.80	1.38	1.09	6.51	10.94	2.74	0.02	*
Speech rate	-7.86	1.13	-10.08	-5.64	488.51	-6.94	0.00	*
Voiced $\times$ /o/	1.91	4.88	-7.65	11.47	11.80	0.39	0.70	
Voiced $\times$ /u/	10.88	4.79	1.50	20.27	10.97	2.27	0.04	*
Voiced $\times$ Polish	2.30	6.07	-9.59	14.19	19.83	0.38	0.71	
/o/ $\times$ Polish	-1.04	4.67	-10.19	8.10	9.94	-0.22	0.83	
/u/ $\times$ Polish	6.94	4.68	-2.24	16.12	10.09	1.48	0.17	
Voiced $\times$ /o/ $\times$ Polish	1.36	6.84	-12.04	14.77	11.44	0.20	0.85	
Voiced $\times$ /u/ $\times$ Polish	-3.08	6.77	-16.35	10.20	11.01	-0.45	0.66	



**Table 4:** Summary of the linear mixed-effects model for testing the correlation between vowel and closure duration (see Section 3.3).

Predictor	Estimate	SE	CI low	CI up	df	t-value	p-value	< $\alpha$
Intercept	123.62	6.76	110.36	136.87	56.24	18.27	0.00	*
Closure dur.	-0.19	0.06	-0.32	-0.06	816.53	-2.93	0.00	*
Vowel = /o/	-4.54	6.31	-16.90	7.82	127.47	-0.72	0.47	
Vowel = /u/	-12.47	6.40	-25.00	0.07	134.65	-1.95	0.05	
Speech rate	-5.16	4.28	-13.55	3.23	827.04	-1.21	0.23	
Closure $\times$ /o/	-0.06	0.08	-0.22	0.10	829.38	-0.71	0.48	
Closure $\times$ /u/	-0.23	0.08	-0.39	-0.07	831.49	-2.82	0.00	*
C2 closure $\times$ sp. rate	-0.20	0.06	-0.32	-0.08	826.97	-3.18	0.00	*
/o/ $\times$ sp. rate	-3.75	5.19	-13.92	6.42	819.79	-0.72	0.47	
/u/ $\times$ sp. rate	-10.13	5.50	-20.91	0.64	822.55	-1.84	0.07	
Closure $\times$ /o/ $\times$ sp. rate	0.09	0.07	-0.06	0.23	820.74	1.17	0.24	
Closure $\times$ /u/ $\times$ sp. rate	0.17	0.08	0.01	0.32	823.88	2.14	0.03	*

**Table 5:** Participants' sociolinguistic information. The column 'Spent most time in' gives the city in which the participant spent most of their life. The last column ('> 6 mo') indicates whether the participant has spent more than 6 months abroad.

ID	Age	Sex	Native L	Other Ls	City of birth	Spent most time in	> 6 mo
IT01	29	Male	Italian	English, Spanish	Verbania	Verbania	Yes
IT02	26	Male	Italian	Friulian, English, Ladin-Venetan	Udine	Tricesimo	Yes
IT03	28	Female	Italian	English, German	Verbania	Verbania	No
IT04	54	Female	Italian	Calabrese	Verbania	Verbania	No
IT05	28	Female	Italian	English	Verbania	Verbania	No
IT09	35	Female	Italian	English	Vignola	Vignola	Yes
IT11	24	Male	Italian	English	Monza	Monza	Yes
IT12	26	Male	Italian	English	Rome	Rome	Yes
IT13	20	Female	Italian	English, French, Arabic, Farsi	Ancona	Chiaravalle	Yes
IT14	32	Male	Italian	English, Spanish	Frosinone	Frosinone	Yes
PL02	32	Female	Polish	English, Norwegian, French, German, Dutch	Koło	Poznań	Yes
PL03	26	Male	Polish	Russian, English, French, German	Nowa Sol	Poznań	Yes
PL04	34	Female	Polish	Spanish, English, French	Warsaw	Warsaw	No
PL05	42	Male	Polish	English, French	Przasnysz	Warsaw	No
PL06	33	Male	Polish	English	Zgierz	Zgierz	Yes
PL07	32	Female	Polish	English, Russian	Bielsk Podlaski	Bielsk Podlaski	Yes

**Table 6:** Target words. Asterisks indicate real words .

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

## References

- Ambridge, Ben. 2018. Against stored abstractions: A radical exemplar model of language acquisition. Pre-print available at PsyArXiv. <https://doi.org/10.31234/osf.io/gy3ah>.
- Ananthapadmanabha, T. V., A. P. Prathosh & A. G. Ramakrishnan. 2014. Detection of the closure-burst transitions of stops and affricates in continuous speech using the plosion index. *The Journal of the Acoustical Society of America* 135(1). 460–471. <https://doi.org/10.1121/1.4836055>.
- Antoniou, Mark, Catherine T. Best, Michael D. Tyler & Christian Kroos. 2010. Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics* 38(4). 640–653. <https://doi.org/10.1016/j.wocn.2010.09.005>.
- Articulate Instruments Ltd™. 2011. Articulate Assistant Advanced user guide. Version 2.16.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Beckman, Jill, Michael Jessen & Catherine Ringen. 2013. Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics* 49(02). 259–284. <https://doi.org/10.1017/S0022226712000424>.
- Beguš, Gašper. 2017. Effects of ejective stops on preceding vowel duration. *The Journal of the Acoustical Society of America* 142(4). 2168–2184. <https://doi.org/10.1121/1.5007728>.
- Belasco, Simon. 1953. The influence of force of articulation of consonants on vowel duration. *The Journal of the Acoustical Society of America* 25(5). 1015–1016.
- 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 lin-

- guistics: A position statement on data citation and attribution in our field. *Linguistics* 56(1). 1–18. <https://doi.org/10.1515/ling-2017-0032>.
- Bigi, Brigitte. 2015. SPPAS - Multi-lingual approaches to the automatic annotation of speech. *The Phonetician* 111–112. 54–69.
- Bonett, Douglas G. 2006. Confidence interval for a coefficient of quartile variation. *Computational Statistics & Data Analysis* 50(11). 2953–2957. <https://doi.org/10.1016/j.csda.2005.05.007>.
- Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22(3). 129–159. <https://doi.org/10.1159/000259312>.
- 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. <https://doi.org/10.31234/osf.io/cfzyx>.
- Cyran, Eugeniusz. 2011. Laryngeal realism and laryngeal relativism: Two voicing systems in Polish? *Studies in Polish Linguistics* 6(1). 45–80.
- Cysouw, Michael & Jeff Good. 2013. Languoid, doculect, and glossonym: Formalizing the notion ‘language’. *Language Documentation & Conservation* 7. 331–359. <https://doi.org/10.125/4606>.
- Dale, Rick, Eric Dietrich & Anthony Chemero. 2009. Explanatory pluralism in cognitive science. *Cognitive science* 33(5). 739–742. <https://doi.org/10.1111/j.1551-6709.2009.01042.x>.
- Davis, Stuart & W. Van Summers. 1989. Vowel length and closure duration in word-medial VC sequences. *Journal of Phonetics* 17. 339–353. <https://doi.org/10.1121/1.2026892>.
- de Jong, Kenneth. 1991. An articulatory study of consonant-induced vowel duration changes in English. *Phonetica* 48(1). 1–17. <https://doi.org/10.1121/1.2028316>.
- de Jong, Kenneth. 2004. Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics* 32(4). 493–516. <https://doi.org/10.1016/j.wocn.2004.05.002>.
- Durvasula, Karthik & Qian Luo. 2012. Voicing, aspiration, and vowel duration in Hindi. *Proceedings of Meetings on Acoustics* 18. 1–10. <https://doi.org/10.1121/1.4895027>.
- Esposito, Anna. 2002. On vowel height and consonantal voicing effects: Data from Italian. *Phonetica* 59(4). 197–231. <https://doi.org/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. [https://doi.org/10.1016/0167-6393\(86\)90027-0](https://doi.org/10.1016/0167-6393(86)90027-0).

- 868 Ferrero, Franco E., Emanuela Magno Caldognetto, Kiryaki Vaggas &  
869 C Lavagnoli. 1978. Some acoustic characteristics of Italian vowels. *Jour-*  
870 *nal of Italian Linguistics Amsterdam* 3(1). 87–94.
- 871 Fowler, Carol A. 1992. Vowel duration and closure duration in voiced and  
872 unvoiced stops: There are no contrast effects here. *Journal of Phonetics*  
873 20(1). 143–165.
- 874 Fox, John. 2003. Effect displays in R for generalised linear models. *Journal*  
875 *of Statistical Software* 8(15). 1–27. [https://doi.org/10.18637/jss.v008.](https://doi.org/10.18637/jss.v008.i15)  
876 [i15](https://doi.org/10.18637/jss.v008.i15).
- 877 Gelman, Andrew & Eric Loken. 2013. The garden of forking paths:  
878 Why multiple comparisons can be a problem, even when there is no  
879 “fishing expedition” or “p-hacking” and the research hypothesis was  
880 posited ahead of time. Department of Statistics, Columbia Univer-  
881 sity, [http://www.stat.columbia.edu/~gelman/research/unpublished/](http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf)  
882 [p\\_hacking.pdf](http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf).
- 883 Gussmann, Edmund. 2007. *The phonology of Polish*. Oxford University Press.
- 884 Hajek, John & Mary Stevens. 2008. Vowel duration, compression and  
885 lengthening in stressed syllables in Central and Southern varieties of  
886 standard Italian. In *Proceedings of the 9th annual conference of the inter-*  
887 *national speech communication association*. 516–519.
- 888 Halle, M., K. N. Stevens & A. V. Oppenheim. 1967. On the mechanism of  
889 glottal vibration for vowels and consonants. In *Quarterly progress report*,  
890 vol. 85. 267–277. <http://hdl.handle.net/1721.1/55578>.
- 891 Halle, Morris & Kenneth Stevens. 1967. Mechanism of glottal vibration for  
892 vowels and consonants. *The Journal of the Acoustical Society of America*  
893 41(6). 1613–1613. <https://doi.org/10.1121/1.2143736>.
- 894 Heffner, R.-M.S. 1937. Notes on the length of vowels. *American Speech* 12.  
895 128–134. <https://doi.org/10.2307/452621>.
- 896 House, Arthur S. & Grant Fairbanks. 1953. The influence of consonant  
897 environment upon the secondary acoustical characteristics of vowels.  
898 *The Journal of the Acoustical Society of America* 25(1). 105–113. <https://doi.org/10.1121/1.1906982>.
- 899 Hualde, José Ignacio & Marianna Nadeu. 2011. Lenition and phonemic  
900 overlap in Rome Italian. *Phonetica* 68(4). 215–242. [https://doi.org/10.](https://doi.org/10.1159/000334303)  
901 [1159/000334303](https://doi.org/10.1159/000334303).
- 902 Hussein, Lutfi. 1994. *Voicing-dependent vowel duration in Standard Arabic*  
903 *and its acquisition by adult American students*: The Ohio State University  
904 dissertation.
- 905 Huszthy, Bálint. 2016. Italian as a voice language without voice assimila-  
906 tion. In *Proceedings of ConSOLE XXIV*. 428–452.
- 907

- 908 Jacewicz, Ewa, Robert Allen Fox & Samantha Lyle. 2009. Variation in stop  
909 consonant voicing in two regional varieties of American English. *Journal*  
910 *of the International Phonetic Association* 39(3). 313–334. [https://doi.org/](https://doi.org/10.1017/S0025100309990156)  
911 [10.1017/S0025100309990156](https://doi.org/10.1017/S0025100309990156).
- 912 Jarosz, Andrew F & Jennifer Wiley. 2014. What are the odds? A practical  
913 guide to computing and reporting Bayes factors. *The Journal of Problem*  
914 *Solving* 7(1). 2–9. <https://doi.org/10.7771/1932-6246.1167>.
- 915 Jassem, Wiktor & Lutoslaw Richter. 1989. Neutralization of voicing in  
916 Polish obstruents. *Journal of Phonetics* 17(4). 317–325.
- 917 Javkin, Hector R. 1976. The perceptual basis of vowel duration differences  
918 associated with the voiced/voiceless distinction. *Report of the Phonology*  
919 *Laboratory, UC Berkeley* 1. 78–92.
- 920 Johnson, Keith. 1997. Speech perception without speaker normalization:  
921 An exemplar model. In Keith Johnson & John W. Mullenix (eds.), *Talker*  
922 *variability in speech processing*, 145–165. San Diego, California: Academic  
923 Press.
- 924 Keating, Patricia A. 1984. Universal phonetics and the organization of gram-  
925 mars. *UCLA Working Papers in Phonetics* 59. [https://escholarship.org/](https://escholarship.org/uc/item/2497n8jq)  
926 [uc/item/2497n8jq](https://escholarship.org/uc/item/2497n8jq).
- 927 Kerr, Norbert L. 1998. HARKing: Hypothesizing after the results are known.  
928 *Personality and Social Psychology Review* 2(3). 196–217. [https://doi.org/](https://doi.org/10.1207/s15327957pspr0203_4)  
929 [10.1207/s15327957pspr0203\\_4](https://doi.org/10.1207/s15327957pspr0203_4).
- 930 Kingston, John & Randy L. Diehl. 1994. Phonetic knowledge. *Language*  
931 419–454. <https://doi.org/10.2307/416481>.
- 932 Kirby, James & Morgan Sonderegger. 2018. Mixed-effects design analysis  
933 for experimental phonetics. *Journal of Phonetics* 70. 70–85. [https://doi.](https://doi.org/10.1016/j.wocn.2018.05.005)  
934 [org/10.1016/j.wocn.2018.05.005](https://doi.org/10.1016/j.wocn.2018.05.005).
- 935 Kirby, James P. & D. Robert Ladd. 2016. Effects of obstruent voicing on  
936 vowel F0: Evidence from “true voicing” languages. *The Journal of the*  
937 *Acoustical Society of America* 140(4). 2400–2411. [https://doi.org/10.](https://doi.org/10.1121/1.4962445)  
938 [1121/1.4962445](https://doi.org/10.1121/1.4962445).
- 939 Klatt, Dennis H. 1973. Interaction between two factors that influence vowel  
940 duration. *The Journal of the Acoustical Society of America* 54(4). 1102–  
941 1104. <https://doi.org/10.1121/1.1914322>.
- 942 Kluender, Keith R., Randy L. Diehl & Beverly A. Wright. 1988. Vowel-  
943 length differences before voiced and voiceless consonants: An auditory  
944 explanation. *Journal of Phonetics* 16. 153–169.
- 945 Ko, Eon-Suk. 2018. Asymmetric effects of speaking rate on the  
946 vowel/consonant ratio conditioned by coda voicing in English. *Phonetics*  
947 *and Speech Sciences* 10(2). 45–50. <https://doi.org/10.13064/KSSS.2018>.

10.2.045.

- 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). <https://doi.org/10.18637/jss.v082.i13>.
- Laeuffer, Christiane. 1992. Patterns of voicing-conditioned vowel duration in French and English. *Journal of Phonetics* 20(4). 411–440.
- 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. <https://doi.org/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. <https://doi.org/10.1121/1.1974906>.
- Lehiste, Ilse. 1970b. Temporal organization of spoken language. In *OSU Working Papers in Linguistics*, vol. 4. 96–114. [https://linguistics.osu.edu/sites/linguistics.osu.edu/files/workingpapers/osu\\_wpl\\_04.pdf](https://linguistics.osu.edu/sites/linguistics.osu.edu/files/workingpapers/osu_wpl_04.pdf).
- Lindblom, Björn. 1967. Vowel duration and a model of lip mandible coordination. *Speech Transmission Laboratory Quarterly Progress Status Report* 4. 1–29. [http://www.speech.kth.se/prod/publications/files/qpsr/1967/1967\\_8\\_4\\_001-029.pdf](http://www.speech.kth.se/prod/publications/files/qpsr/1967/1967_8_4_001-029.pdf).
- Lisker, Leigh. 1957. Closure duration and the intervocalic voiced-voiceless distinction in English. *Language* 33(1). 42–49. <https://doi.org/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. <https://doi.org/10.1121/1.392651>.
- Luke, Steven G. 2017. Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods* 49(4). 1494–1502. <https://doi.org/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. <https://escholarship.org/uc/item/31f5j8m7>.



- 987 Magno Caldognetto, Emanuela, Franco Ferrero, Kyriaki Vaggas & Maria  
988 Bagno. 1979. Indici acustici e indici percettivi nel riconoscimento  
989 dei suoni linguistici (con applicazione alle consonanti occlusive  
990 dell'italiano). *Acta Phoniatica Latina* 2. 219–246.
- 991 Malisz, Zofia & Katarzyna Klessa. 2008. A preliminary study of temporal  
992 adaptation in Polish VC groups. In *Proceedings of speech prosody*. 383–  
993 386.
- 994 McElreath, Richard. 2015. *Statistical rethinking: A bayesian course with ex-*  
995 *amples in R and Stan*. CRC Press.
- 996 Meyer, Ernst Alfred. 1904. Zur vokaldauer im deutschen. In *Nordiska studier*  
997 *tillegnade A. Noreen*, 347–356. K.W. Appelbergs Boktryckeri: Uppsala.
- 998 Mitleb, Fares. 1982. Voicing effect on vowel duration is not an absolute  
999 universal. *The Journal of the Acoustical Society of America* 71(S1). S23–  
1000 S23. <https://doi.org/10.1121/1.2019285>.
- 1001 Nicenboim, Bruno, Timo B. Roettger & Shravan Vasishth. 2018. Using meta-  
1002 analysis for evidence synthesis: The case of incomplete neutralization  
1003 in german. *Journal of Phonetics* 70. 39–55. [https://doi.org/10.1016/j.](https://doi.org/10.1016/j.wocn.2018.06.001)  
1004 [wocn.2018.06.001](https://doi.org/10.1016/j.wocn.2018.06.001).
- 1005 Nowak, Pawel. 2006. *Vowel reduction in Polish*: University of California,  
1006 Berkeley dissertation.
- 1007 Ohala, John J. 2011. Accommodation to the aerodynamic voicing constraint  
1008 and its phonological relevance. In *Proceedings of the 17th International*  
1009 *Congress of Phonetic Sciences*. 64–67.
- 1010 Peterson, Gordon E. & Ilse Lehiste. 1960. Duration of syllable nuclei in  
1011 English. *The Journal of the Acoustical Society of America* 32(6). 693–703.  
1012 <https://doi.org/10.1121/1.1908183>.
- 1013 Plug, Leendert & Rachel Smith. 2018. Segments, syllables and speech tempo  
1014 perception. In *Proceedings of the 9th International Conference on Speech*  
1015 *Prosody 2018*. 279–283. [https://doi.org/10.21437/SpeechProsody.](https://doi.org/10.21437/SpeechProsody.2018-57)  
1016 [2018-57](https://doi.org/10.21437/SpeechProsody.2018-57).
- 1017 Port, Robert F. & Jonathan Dalby. 1982. Consonant/vowel ratio as a cue  
1018 for voicing in English. *Perception & Psychophysics* 32(2). 141–152.
- 1019 R Core Team. 2018. R: A language and environment for statistical com-  
1020 puting. R Foundation for Statistical Computing, Vienna, Austria. [https:](https://www.R-project.org)  
1021 [//www.R-project.org](https://www.R-project.org).
- 1022 Raftery, Adrian E. 1995. Bayesian model selection in social research. *Soci-*  
1023 *ological methodology* 111–163. <https://doi.org/10.2307/271063>.
- 1024 Raftery, Adrian E. 1999. Bayes factors and BIC: Comment on “A cri-  
1025 tique of the Bayesian information criterion for model selection”. *So-*  
1026 *ciological Methods & Research* 27(3). 411–427. <https://doi.org/10.1177/>

- 1027 0049124199027003005.
- 1028 Renwick, Margaret & Robert D. Ladd. 2016. Phonetic distinctiveness vs.  
1029 lexical contrastiveness in non-robust phonemic contrasts. *Laboratory*  
1030 *Phonology: Journal of the Association for Laboratory Phonology* 7(1). 1–  
1031 29. <https://doi.org/10.5334/labphon.17>.
- 1032 Roettger, Timo B. 2018. Researcher degrees of freedom in phonetic sciences.  
1033 Pre-print available at PsyArXiv. <https://doi.org/10.31234/osf.io/fp4jr>.
- 1034 Roettger, Timo B. 2019. Researcher degrees of freedom in phonetic sciences.  
1035 *Laboratory Phonology: Journal of the Association for Laboratory Phonology*  
1036 10(1). 1–27. <https://doi.org/10.5334/labphon.147>.
- 1037 Rothenberg, Martin. 1967. *The breath-stream dynamics of simple-released-*  
1038 *plosive production*, vol. 6. Basel: Biblioteca Phonetica.
- 1039 Sanker, Chelsea. 2018. Effects of laryngeal features on vowel duration:  
1040 implications for Winter's Law. *Papers in Historical Phonology* 3. 180–205.  
1041 <https://doi.org/10.2218/pihph.3.2018.2898>.
- 1042 Schwartz, Geoffrey. 2016. On the evolution of prosodic boundaries–  
1043 parameter settings for Polish and English. *Lingua* 171. 37–73. <https://doi.org/10.1016/j.lingua.2015.11.005>.
- 1044 Schwartz, Geoffrey & Daria Arndt. 2018. Laryngeal Realism vs. Modulation  
1045 theory – evidence from VOT discrimination in Polish. *Language Sciences*  
1046 69. 98–112. <https://doi.org/10.1016/j.langsci.2018.07.001>.
- 1047 Schwartz, Geoffrey, Anna Balas & Arkadiusz Rojczyk. 2015. Phonological  
1048 factors affecting L1 phonetic realization of proficient Polish users of En-  
1049 glish. *Research in Language* 13(2). 181–198. <https://doi.org/10.1515/rela-2015-0014>.
- 1050  
1051
- 1052 Sharf, Donald J. 1962. Duration of post-stress intervocalic stops and pre-  
1053 ceding vowels. *Language and speech* 5(1). 26–30.
- 1054 Sharf, Donald J. 1964. Vowel duration in whispered and in normal speech.  
1055 *Language and speech* 7(2). 89–97.
- 1056 Slis, Iman Hans & Antonie Cohen. 1969a. On the complex regulating the  
1057 voiced-voiceless distinction I. *Language and speech* 12(2). 80–102. <https://doi.org/10.1177/002383096901200202>.
- 1058  
1059 Slis, Iman Hans & Antonie Cohen. 1969b. On the complex regulating the  
1060 voiced-voiceless distinction II. *Language and speech* 12(3). 137–155.  
1061 <https://doi.org/10.1177/002383096901200301>.
- 1062 Slowiaczek, Louisa M. & Daniel A. Dinnsen. 1985. On the neutralizing status  
1063 of Polish word-final devoicing. *Journal of Phonetics* 13(3). 325–341.
- 1064 Sóskuthy, Márton. 2013. *Phonetic biases and systemic effects in the actuation*  
1065 *of sound change*: University of Edinburgh dissertation.



- 1066 Sós-kuthy, Márton, Paul Foulkes, Vincent Hughes & Bill Haddican. 2018.  
1067 Changing words and sounds: the roles of different cognitive units in  
1068 sound change. *Topics in cognitive science* 1–16. [https://doi.org/10.1111/](https://doi.org/10.1111/tops.12346)  
1069 [tops.12346](https://doi.org/10.1111/tops.12346).
- 1070 Strycharczuk, Patrycja. 2012. Sonorant transparency and the complexity of  
1071 voicing in Polish. *Journal of Phonetics* 40(5). 655–671. [https://doi.org/](https://doi.org/10.1016/j.wocn.2012.05.006)  
1072 [10.1016/j.wocn.2012.05.006](https://doi.org/10.1016/j.wocn.2012.05.006).
- 1073 Tilsen, Sam. 2013. A dynamical model of hierarchical selection and coor-  
1074 dination in speech planning. *PLoS ONE* 8(4). e62800. [https://doi.org/](https://doi.org/10.1371/journal.pone.0062800)  
1075 [10.1371/journal.pone.0062800](https://doi.org/10.1371/journal.pone.0062800).
- 1076 Tilsen, Sam. 2016. Selection and coordination: The articulatory basis for  
1077 the emergence of phonological structure. *Journal of Phonetics* 55. 53–77.  
1078 <https://doi.org/10.1016/j.wocn.2015.11.005>.
- 1079 Todd, Simon, Janet B. Pierrehumbert & Jennifer Hay. 2018. Word frequency  
1080 effects in sound change as a consequence of perceptual asymmetries: An  
1081 exemplar-based model. Submitted.
- 1082 Van Summers, W. 1987. Effects of stress and final-consonant voicing on  
1083 vowel production: Articulatory and acoustic analyses. *The Journal of the*  
1084 *Acoustical Society of America* 82(3). 847–863. [https://doi.org/10.1121/](https://doi.org/10.1121/1.395284)  
1085 [1.395284](https://doi.org/10.1121/1.395284).
- 1086 Vazquez-Alvarez, Yolanda & Nigel Hewlett. 2007. The ‘trough effect’: an  
1087 ultrasound study. *Phonetica* 64(2-3). 105–121. [https://doi.org/10.1159/](https://doi.org/10.1159/000107912)  
1088 [000107912](https://doi.org/10.1159/000107912).
- 1089 Wagenmakers, Eric-Jan. 2007. A practical solution to the pervasive prob-  
1090 lems of *p* values. *Psychonomic bulletin & review* 14(5). 779–804. <https://doi.org/10.3758/BF03194105>.  
1091 <https://doi.org/10.3758/BF03194105>.
- 1092 Waniek-Klimczak, Ewa. 2011. Aspiration in Polish: A sound change in  
1093 progress? In Mirosław Pawlak & Jakub Bielak (eds.), *New perspec-*  
1094 *tives in language, discourse and translation studies*, 3–11. Springer. [https://doi.org/10.1007/978-3-642-20083-0\\_1](https://doi.org/10.1007/978-3-642-20083-0_1).  
1095 [https://doi.org/10.1007/978-3-642-20083-0\\_1](https://doi.org/10.1007/978-3-642-20083-0_1).
- 1096 Warren, Willis & Adam Jacks. 2005. Lip and jaw closing gesture durations  
1097 in syllable final voiced and voiceless stops. *The Journal of the Acousti-*  
1098 *cal Society of America* 117(4). 2618–2618. [https://doi.org/10.1121/1.](https://doi.org/10.1121/1.4778168)  
1099 [4778168](https://doi.org/10.1121/1.4778168).
- 1100 Wells, John C. 1990. Syllabification and allophony. *Studies in the pronuncia-*  
1101 *tion of English: A commemorative volume in honour of A. C. Gimson* 76–86.
- 1102 Westbury, John R. 1983. Enlargement of the supraglottal cavity and its  
1103 relation to stop consonant voicing. *The Journal of the Acoustical Society*  
1104 *of America* 73(4). 1322–1336.

- 1105 Wickham, Hadley. 2017. tidyverse: Easily install and load the ‘Tidy-  
1106 verse’. R package version 1.2.1. [https://CRAN.R-project.org/package=](https://CRAN.R-project.org/package=tidyverse)  
1107 [tidyverse](https://CRAN.R-project.org/package=tidyverse).
- 1108 Winter, Bodo & Timo B. Roettger. 2011. The nature of incomplete neu-  
1109 tralization in German: Implications for Laboratory Phonology. *Grazer*  
1110 *Linguistische Studien* 76. 55–74.
- 1111 Zuur, Alain F., Elena N. Ieno & Chris S. Elphick. 2010. A protocol for data  
1112 exploration to avoid common statistical problems. *Methods in ecology*  
1113 *and evolution* 1(1). 3–14. [https://doi.org/10.1111/j.2041-210X.2009.](https://doi.org/10.1111/j.2041-210X.2009.00001.x)  
1114 [00001.x](https://doi.org/10.1111/j.2041-210X.2009.00001.x).