SOCIOPHONETICS AND STOPS

Eleanor Chodroff and Paul Foulkes
University of York

1. Introduction

All languages have stop consonants. The voiceless stops /p t k/ occur in around 98% of languages (Moran and McCloy 2019), and "every known language has sounds similar to two of these three" (Ladefoged 2005: 156). Stops are also acquired early by children learning almost any language (Vihman 1996: 117; McLeod and Crowe 2018).

While the prevalence and early acquisition might suggest some degree of simplicity, from a phonetic perspective stops are complex. Consider the articulatory actions involved in generating a plosive. First, an active articulator is brought into contact with a passive articulator. The closure is held while pulmonic air pressure builds up behind it. Plosion occurs as the constriction is released, and the articulators then continue to move in order to produce the subsequent sound (or to return to rest). All of those supralaryngeal actions must also be coordinated in time with laryngeal action, to ensure, for example, that vocal fold vibration starts and ceases appropriately. The sequence of articulatory changes thus generates a complex series of aerodynamic and acoustic changes. Airflow is first channelled through an increasingly narrow cavity as the closure is formed, then air pressure builds during the hold phase before being released abruptly and generating the sudden stop 'burst' as high pressure air escapes the oral tract. Following the release there is a sudden drop in air pressure, before airflow continues through a vocal tract that is changing for the production of subsequent sounds. In acoustic terms we see a series of events including (depending on the precise sequence of sounds) formant transitions in a preceding voiced sound, acoustic silence during a voiceless closure or low amplitude periodic energy during a voiced one, transient(s) at the point of release, aperiodic noise as air escapes the oral tract, further acoustic transitions in any subsequent sound, and ongoing changes in phonation throughout the sequence.

This physical complexity presents a wide range of opportunities for variation in the phonetic qualities of stops. Variation can be observed in the closing phase, the hold phase, the release phase, and/or in the timing relationship between supralaryngeal and laryngeal actions. From a cross-linguistic perspective, Ladefoged and Maddieson (1996: 47-48) note that stops vary in their precise place of articulation, glottal state, air stream mechanism, activity at onset and offset (e.g. pre-aspiration and post-aspiration), length, and strength. Variation can be gradient (e.g. in the duration of voice onset time) or categorical (e.g. in terms of phonological units being deleted, inserted or substituted). Similar patterns of variation can, in principle, be observed within and between speakers of a given language. Variation may partially reflect universal constraints of biology and physics, but it may also reflect systematic regional and social differences that convey indexical meaning.

In this chapter we review sources and patterns of variation in stops. We focus mainly on speech production, but also briefly discuss perceptual studies. In section 2 we illustrate variation at the phonetic and phonological levels, before turning to review examples of systematic regional and social variation in section 3. Finally, in section 4 we describe a case study conducted on variation across speaker sex and age in Derby English stop realisation as measured through voice onset time (VOT) and spectral centre of gravity (COG). In particular, we focus on the locus of variation for social variables in the process of phonetic realisation: do social variables influence the phonetic realisation of individual speech sounds as independent units, or do social variables associate with the phonetic realisation of shared phonological distinctive features? The case study exemplifies a corpus phonetics approach to sociophonetic variation. Through this approach we uncover systematic variation in word-initial stops relating linguistic factors to social factors of sex and age.

2. Phonetic and phonological variation in stops

As noted in the introduction, stops are dynamic in terms of their articulatory actions and their aerodynamic and acoustic consequences. There are at least three distinct phases of the active oral articulator, which are coupled with a laryngeal gesture. In the following three sections, we outline phonetic variation in terms of these three phases: closing, hold, and release. Section 2.4 discusses laryngeal coupling and voice onset time, and finally section 2.5 illustrates common phonological alternations observed with stops. To some extent the examples of variation reflect language-specific patterns rather than social ones, but in principle the same factors may lead to variation between social or regional groups within a language, or between individual speakers (see further section 3).

2.1. Closing phase

The closing gesture for a stop may be executed at varying speed, and with varying coordination relative to laryngeal activity. In some circumstances, for example if closure is relatively slow and phonation ceases prior to full closure, there may be a period in which a relatively narrow channel in the oral tract coincides with high airflow. These conditions may generate aperiodic noise at the point of constriction (i.e. pre-affrication) and/or at the glottis or throughout the vocal tract (i.e. pre-aspiration). The presence or absence of pre-aspiration can signal a phonological contrast, as in Gaelic and Icelandic. It is also common but non-contrastive in other Nordic languages (Helgason 2002, Hejná 2015, Clayton 2017), and reported as a social and regional feature in some accents of English (see further section 3.1). In Greek there is debate on how the closing phase is coordinated with velic action. Voiced stops in some dialects are preceded by a nasal element, leading some scholars to treat them as prenasalised, while others treat them as sequences of nasal+homorganic stop. However, recent studies suggest that the overall duration is not increased if there is a nasal element, supporting the prenasalisation interpretation (Arvaniti 2007).

Rate of closure varies by place of articulation. The velocity of the tongue body is typically slower than that of the tongue tip; the following vowel also influences this rate, with low vowels corresponding to faster closing rates than high vowels, at least for tongue tip and tongue body closures (Löfqvist and Gracco 2002). Moreover, alveolars and velars are often influenced by adjacent sounds (front vowels leading to fronter stop articulation, and back vowels conversely leading to retraction).

Another common source of variation is found in the magnitude of the closure. While stops canonically involve complete closure, in some accents and some circumstances (e.g. particular phonological contexts, or in rapid speech) the full closure is not made. Instead speakers generate narrow but incomplete constrictions, thus producing phonetic fricatives or even approximants. Such events are often described as one type of lenition or weakening (e.g. Lass 1984, Kirchner 2001, Gurevich 2004). Stop lenition, especially in intervocalic and non-initial positions, is reported in many languages including Spanish (Hualde 2005), Shilluk (spoken in Sudan and South Sudan; Remijsen et al. 2011), and Paya Kuna (spoken in Panama and Colombia; Pike et al. 1986). These lenition patterns frequently target a full laryngeal series (e.g. the voiced stop series in Spanish). Nevertheless, the propensity for lenition varies across languages, segments and phonological contexts. Cohen Priva (2017) identifies informativity as one determining factor underlying this variability: languages are more tolerant of lenition where the segments concerned offer relatively low information content (defined as the mean predictability of segments in context, rather than raw frequency). Such work is facilitated by corpus-based methods which allow access to the large data sets required to assess factors such as predictability and frequency.

2.2. Hold phase

Once a closure is formed it may be held for varying durations. The duration of this closure, or hold, is frequently correlated with laryngeal activity. Specifically, there seems to be a universal pattern in which hold duration and voice onset time (VOT) are in an inverse correlation with respect to place of articulation (see also section 2.4). Labials have the longest hold duration and shortest VOT, while velars have short hold and long VOT (Byrd 1993; Maddieson 1997; Yao 2007; Phillips and Tucker 2020). In a study of American English, Yao (2007) observed some variation between individual speakers in closure duration, although [p^h] had the longest hold duration for all speakers. The correlation between hold duration and VOT likely reflects the fact that phonation can initiate only when supraglottal pressure is lower than subglottal pressure. This has a direct relationship with place of articulation, as the constriction location results in different intraoral volumes. Once the closure is made, supraglottal pressure immediately begins to increase as air accumulates behind the closure. Supraglottal pressure may come to equal that of subglottal pressure, thus rendering vocal fold vibration impossible without some sort of articulatory adjustment, such as lowering of the larynx, expanding the cheeks, or releasing air by lowering the velum (Westbury 1983). A larger intraoral volume corresponds to a lower baseline supraglottal pressure. In voiced stops this allows for longer vocal fold vibration during the hold, resulting in a more negative (but greater magnitude) VOT for [b] than [g]. For instance, in Lisker and Abramson's (1964) data for Puerto Rico Spanish, mean VOT for [b] was -138 ms while that for [g] was -108 ms. In voiceless stops, in which phonation must *follow* the hold phase, a larger intraoral volume requires more time for a rise in supraglottal pressure, resulting in a long hold duration and correspondingly shorter VOT ([p] < [k]; see further section 2.4).

Davidson (2016) offers a detailed study of stops in American English, showing that phonation from preceding sonorants frequently 'bleeds' into the hold phase of stops (~80% of all cases), but varies in respect of word and phrase position, stress, and the type of preceding segment. A case of 'bleeding' is voicing that continues from a preceding vowel into the hold phase, but decreases in intensity and ceases in the final third of the interval. Because supraglottal pressure quickly rises in stops, partial devoicing of phonologically voiced stops is therefore common. In addition, where a phonologically voiced series is cued by zero or short VOT, such as in standard English, many speakers in fact also occasionally or regularly use pre-voicing (Lisker and Abramson 1964, Docherty 1992). Where pre-voicing is used, the phonation is often extinguished prior to the release. As described above, vocal fold vibration becomes impossible if intraoral air pressure rises to equalise or exceed subglottal pressure.

In some languages duration plays an additional phonological role, where geminates (sounds with a long hold) contrast with singleton consonants (short hold).The precise timing relationship between geminate and singleton differs across and within languages. The ratio between geminate and singleton duration varies from less than 2:1 to as much as 3:1 (Khattab and Al-Tamimi 2014; see also Ham 2001, Aoyama and Reid 2006, and papers in Kubozono

2017). Ham (2001) hypothesises that the timing ratio is larger for mora-timed than syllable-timed languages. Cues other than duration may also be used to signal the contrast. For example, vowels tend to be shorter before geminates than before singletons (Al-Tamimi and Khattab 2015). Al-Tamimi and Khattab (2018) provide a detailed discussion of voicing in the context of the singleton-geminate contrast in Lebanese Arabic, drawing on a set of 19 acoustic measures. Ultrasound studies indicate that geminates involve greater tongue raising than singletons (Percival et al. 2018 on Eastern Oromo, Percival et al. 2020 on Hungarian). Closure duration is also relevant for the contrast between a full plosive and a tap: in intervocalic position, it appears common for alveolar stops to reduce from full plosives to taps, as in Shoshone and varieties of English spoken in North America and Australasia (Gurevich 2004).

Variation in laryngeal action is also evident in many languages that have glottal 'stops'. Although the IPA classification of [ʔ] implies that it is parallel to oral stops in terms of its articulation, it is often in fact lenited to a period of creaky voicing rather than executed as a full glottal closure+hold+release sequence. This is true both for languages with a contrastive /ʔ/ such as Hawai'ian (Davidson 2021) and also for languages like English and German where it is allophonic (Docherty and Foulkes 1999, Kohler 1994). A full glottal closure appears most likely in word-initial position.

2.3. Release phase

Stop releases (also referred to as bursts) may vary in intensity and frequency characteristics. Though Lisker (1986) identified burst intensity as a potential correlate of the voicing contrast, with voiceless stops tending to have louder bursts than voiced, research findings have been somewhat mixed. Greater intensity of voiceless than voiced stops has been found in American English (Zue 1976), Canadian English and Canadian French (Sundara et al. 2006), European French (Sundara 2005), and P'urépecha (Zerbe 2013: fortis > lenis). However, the opposite pattern has been observed in Georgian (Vicenik 2008: voiced > voiceless) and Itunyoso Trique (DiCanio 2012: lenis > fortis). Mixed outcomes have been observed depending on the place of articulation in Danish (Fischer-Jørgensen 1954). In Korean, aspirated stops have more intense bursts than lenis and fortis stops (see further section 2.4 below; Cho et al. 2002).

Spectral shape of the acoustic frequency spectrum is commonly measured for understanding aspects of the release burst, particularly the stop place of articulation (Blumstein and Stevens 1979, 1980; Stevens and Blumstein 1981; Kewley-Port 1983; Bonneau et al. 1996). Labial stops have greater energy at lower frequencies (diffuse-falling spectrum), whereas coronal stops have greater at higher frequencies (diffuse-rising spectrum), and dorsal stops have spectral energy concentrated in the mid-frequency range (compact spectrum). In contrast to the large body of work on place, only a few studies have investigated spectral shape as a cue to the voicing contrast. Several studies have also reported greater energy in the higher frequencies for word-initial voiceless stops relative to voiced stops, at least for the labial and coronal places of articulation (American English: Halle et al. 1957; Zue 1976; Parikh and Loizou 2005; Chodroff and Wilson 2014; Canadian English: Sundara 2005; British English: Kirkham 2011; Dutch: van Alphen and Smits 2004; German: Harrington 2010: p. 192; Georgian: Vicenik 2010). These tendencies in the frequency spectrum are frequently summarised via metrics like the spectral peak or spectral moments (centre of gravity: COG, variance or standard deviation, skewness, and kurtosis).

Mirroring the variation found in the closing phase, it is also possible to observe variation in the release phase in terms of speed of release, and coordination relative to laryngeal activity (see 2.4 below). A relatively slow release may generate aperiodic noise at the point of articulation, i.e. affrication or assibilation. This occurs most frequently before high vowels and /j/, and is reported, for example, in Finnish, Québecois French, Korean, Basque, Polish and West Futuna-Aniwa (Hall and Hamann 2006). Release of the air trapped within the oral tract is usually executed by lowering the active articulator to allow the air to escape forwards, but for some speakers the release may be lateral if a lateral sound follows (cf., for some English speakers, *adder* [adə] versus *addle* [adˡɫ]) or nasal, via lowering the velum (cf. *butter* [bʌtə] versus *button* [bʌtⁿn̩]). Davidson (2011) explores stop releases in American English spontaneous speech, reporting a range of factors that affect whether a coda-final stop is released. These factors include the place of articulation, the manner of the following segment, and the stop location (word-medial, word-final, or phrase-final).

In some cases strong releases are generated by a temporary switch in airstream mechanism from pulmonic to glottal, i.e. to produce ejectives. Ejectives are contrastive in many languages (Vicenik 2010), but occur as allophonic variants of voiceless pulmonic stops in other languages. For instance, they are fairly widespread in utterance-final position for many speakers of German and British English, and occur more for /k/ than /p/ or /t/ (McCarthy and Stuart-Smith 2013, Brandt and Simpson 2021). Kuang (2019) discusses variation in both ejective and implosive stops relating to phonological context in the Mayan language Q'anjob'al. Percival (2019) considers contextual variation for ejectives in the Salish language Hul'q'umi'num', while Brandt and Simpson (2021) report a similar study on Georgian and German.

For some speakers, releases themselves may be complex, articulated not as a single physical opening and acoustic transient, but as a series of them. This phenomenon of multiple bursts is reported for English (e.g. Lavoie 2001, Yao 2007, Parveen and Goberman 2012), Hungarian (Gráczi and Kohári 2014) and Spanish (Lavoie 2001), and appears mostly to affect posterior places of articulation. /k/ in particular often displays two or three transients, though six or even more have been reported (Parveen and Goberman 2012). Various explanations have been offered for why multiple releases occur, including gradual release, additional closures being generated by the Bernouilli effect, and transients reflecting 'pops' as saliva is sucked from the articulators (Foulkes et al. 2010: 65).

Stops exert epiphenomenal effects on following vowels. The f0 at the start of a vowel is higher after a voiceless stop than after a voiced stop (Lee et al. 2020). The f0 difference has come to be the dominant cue over time in some languages (for example Korean; Kang & Guion 2008), leading to the emergence of tonal differences (Hombert et al. 1979).

Formant transitions in vowels provide information about the place of articulation of adjacent stops. Following a labial stop all formants rise at the onset of a following vowel, and a reverse pattern of falling formants at the end of the vowel can be observed in vowel+labial sequences. In alveolars and velars the F2 and F3 patterns vary depending on the specific vowel+consonant sequence (Kewley-Port 1982). However, F3 can be especially useful in separating velars and alveolars (Öhman 1966; Fant 1973; Cassidy and Harrington 1995). After a velar, F3 and F2 at the onset of a vowel often appear to emerge from a single point in the frequency range, giving a visual impression of a 'pinch'.

2.4. VOT and laryngeal timing

Around 60% of languages display a voicing contrast between stops produced at the same place of articulation (as English does with its 'voiceless' series /p t k/ and 'voiced' series /b d g/; Maddieson 2013). While the label 'voicing' implies a difference in phonation, the voicing contrast frequently relates back to the phonological contrast between two (or more) series of stop consonants that differ in their laryngeal realisation. The precise label used to express this phonological contrast varies across researchers and frameworks: these include descriptions such as voiced vs voiceless, [±voice], [±spread glottis], fortis vs lenis. Though typologically rare, some languages also have three- or even four-way laryngeal contrasts. Languages with three-way contrasts include Thai (voiced vs voiceless unaspirated vs voiceless aspirated) and Korean (fortis vs lenis vs aspirated). Languages with four-way contrasts include Hindi and Urdi, among others (voiced unaspirated vs voiced aspirated vs voiceless unaspirated vs voiceless aspirated). These are sometimes phonologically classified as being a two-way voicing contrast with a two-way aspiration contrast (e.g. Schertz and Khan 2020).

In stop production, supralaryngeal actions must be coordinated in time with laryngeal ones to ensure that voicing cues are appropriate for the phonological status of the stop (see also section 2.2). In principle, given their IPA descriptors, this means that voiceless stops such as /p t k/ have no vocal fold vibration during the stop articulation, while voiced stops such as /b d g/ have vocal fold vibration throughout their articulation. However, phonological /b d g/ are often cued by phonation starting more or less simultaneously with the release.

Lisker and Abramson (1964) provide benchmark VOT values to typify three primary laryngeal categories that apply across most languages: -125 to -75 ms for pre-voiced stops (lead VOT: voicing precedes the stop release, and is thus conventionally given a negative value), 0-25 ms for unaspirated stops (short-lag VOT: a short lag exists between stop release and onset of voicing), and 60-100 ms for aspirated stops (long-lag VOT: a long lag exists between stop release and voicing). Note that there are differences across studies in how exactly the boundaries of VOT are identified (Foulkes et al. 2010; Abramson and Whalen 2017). These three modes have largely been corroborated in cross-linguistic studies of VOT (Cho and Ladefoged 1999; Chodroff et al. 2019; Cho, Whalen and Docherty 2019), though it is debatable whether such a rigid demarcation between the three categories is actually meaningful in phonological representations (Docherty 1992; Scobbie 2006). Indeed several languages have what could be considered medium-lag VOT (e.g. Lebanese Arabic: Khattab 2003; Hebrew: Raphael et al. 1995; Japanese: Riney et al. 2007). Theoretical frameworks differ as to what degree the state of the larynx reflects the phonological representation of the stop (e.g. laryngeal realism: Beckman et al. 2013 vs. substance-free phonology: Hale and Reiss 2000).

Classic studies of categorical perception typically use stop VOT as a marquee case. Listeners tend to have sharp categorization boundaries along the VOT continuum corresponding to the laryngeal contrast (Liberman and Abramson 1967). The exact location of the boundary varies, however, by language, place of articulation, vowel context, other acoustic properties of the signal, and even by individual (Stevens and Klatt 1974; Keating et al. 1981). Discrimination is extremely diminished between stops identified with the same label compared to those that cross the VOT boundary. Categorical perception of stop VOT has been found even in one month old infants (Eimas et al. 1971). More recent studies, however, have argued that perception of the laryngeal contrast is still encoded gradiently even though the phonological encoding could be categorical. This has been evidenced through neural event-related potential measures in response to variation in VOT (Blumstein et al. 2005; Toscano et al. 2010), as well as the fact that listeners employ multiple acoustic cues to determine stop identity. A listener's categorization function can change as a result of slight differences in combinations of VOT, following vowel length, and other cues. Moreover, this 'cue weighting' can be language-, context- and talker-specific (e.g. Toscano and McMurray 2010).

In languages with more than a two-way laryngeal contrast, the perception of each phonation type becomes more complicated and can depend on the exact constellation of acoustic correlates that separate the phonation categories. For example, the phonetic realisation of Korean's three-way laryngeal contrast makes use of both VOT and f0. For the fortis–aspirated contrast, VOT serves as the primary cue, whereas for the lenis–aspirated contrast, the relative influence of both VOT and f0 depends on the particular dialect (see also section 3.4). In Hindi and Urdu, listeners make use of multiple cues including the presence and degree of prevoicing, aspiration, and voice quality (murmur or breathy voice; Schertz and Khan 2020). VOT is also a secondary cue that distinguishes emphatic from plain stops: for example, emphatic /ṭ/ has a shorter VOT than the corresponding plain /t/ in Jordanian Arabic (Khattab et al. 2006).

VOT varies substantially not only across languages, but also across individual speakers of a given language. To cite just one example, reported averages for the VOT of /t/ in studies of standard accents of English range from 49 ms (Byrd 1993) to 98 ms (Chodroff and Wilson 2017), with further variation between speakers. VOT variation between stop categories, however, is relatively structured across both languages and individuals. As noted in section 2.2, labial stop VOTs are almost always shorter than corresponding dorsal stop VOTs within the same laryngeal series (e.g. Lisker and Abramson 1964; Klatt 1975; Nearey and Rochet 1994; Chodroff et al. 2019). While the coronal VOT value is frequently intermediate to the labial and dorsal VOT values, considerable variation in its exact ranking has been reported (e.g. Suomi 1980; Gandour et al. 1986; Docherty 1992; Yao 2009; Chodroff and Wilson 2017; Chodroff et al. 2019).

Beyond the ordinal relationship of VOT across place of articulation (/p/ < /k/), languages and speakers also exhibit highly linear VOT relationships. Not only is the VOT of /p/ universally shorter than that of /k/, but the difference between /p/ and /k/ is very predictable regardless of whether the VOT is long or short. Across talkers and languages, the mean VOTs for [p^h], [t^h] and [k^h] are highly correlated with one another: as the mean VOT for [p^h] increases, so does the VOT for [t^h] and [k^h] (Chodroff and Wilson 2017; Chodroff et al. 2019). This particularly holds true of stop series with long-lag or lead VOT, but also to some degree for stops with short-lag VOT. This tight relationship between mean values across individuals and languages likely relates to a relatively uniform timing relationship between the laryngeal activity and the hold phase for each place of articulation (Chodroff and Wilson 2017). Moreover, listeners are sensitive to this relationship across place of articulation: after exposure to a talker producing [p^h], listeners learn not only the characteristic VOT for [p^h], but also that for [k^h] (Theodore and Miller 2010). This has also been observed in phonetic imitation studies in which listeners lengthened VOT not only after exposure to a long VOT [p^h], but also for unheard [k^h] (Nielsen 2011). The degree to which talkers imitate VOT, however, varies across individuals (Yu et al. 2013; Wade et al. 2021).

In addition to phonological voicing and place of articulation, VOT also varies in response to many factors including the height of the following vowel, prosodic position, stress, speech rate, word frequency, and phonological neighbourhood density (Chodroff and Wilson 2017). Vowel height and tenseness act as a source of variability in the VOT of voiceless stops. In general, VOT is longer before high vowels than mid or low vowels (Klatt 1975; Flege et al. 1988; Esposito 2002; Berry and Moyle 2011; but cf. Mortensen and Tøndering 2013); additionally, VOT tends to be longer before tense vowels than lax vowels, for example [i u a] versus [ɪ ʊ ʌ] in English (Port and Rotunno 1979; Weismer 1979; Nearey and Rochet 1994). Nearey and Rochet (1994) indicate considerable individual differences in this effect, but do emphasise a consistent lengthening effect of following [i]. Cho, Jun and Ladefoged (2002) also suggest that a key cue to the Korean lenis/fortis contrast lies in the voice quality of the following vowel. Lenis stops are followed by breathier vowels (measured as H1-H2, i.e. the amplitude difference between the first two harmonics).

Prosodic position contributes substantially to variability in word-initial VOT. Voiceless stops in phrase-initial position have a lengthened VOT relative to other positions in English (Pierrehumbert and Talkin 1992; Choi 2003; Chodroff and Wilson 2017) and Korean (Cho and Keating 2001, 2009). In English, significantly longer VOTs are observed for word-initial stops with a phrasal accent than those without, and the effect is stronger for voiceless than for voiced stops (Cole et al. 2003, 2007). Longer VOTs are also observed for word-initial labial stops with a focal accent than those without (Choi 2003). Cho and McQueen (2005) compared VOT in prosodically strong and weak positions in Dutch and English. For /t/ they found that prosodic strength led to shorter VOT in Dutch and longer VOT in English, thus enhancing the key distinctive feature for the phonological contrast (framed as {−spread glottis} for Dutch and {+spread glottis} for English).

Speech rate also has a strong influence on the VOT of voiceless stops. Faster speaking rates correspond to shorter VOTs in production and perception (e.g. Summerfield 1975; Theodore et al. 2009). This effect has been confirmed with measures of speaking rate that include syllable duration (e.g. Miller et al. 1986; Kessinger and Blumstein 1997, 1998), following vowel duration (e.g. Port and Rotunno 1979; Allen and Miller 1999; Allen et al. 2003; Theodore et al. 2009) and number of syllables per second (Baum and Ryan 1993).

Lexical properties such as frequency and phonological neighbourhood density affect the realisation of VOT. Higher frequency words are associated with shorter VOTs, at least in English word-initial stops (Yao 2009; Chodroff and Wilson 2017, Mielke and Nielsen 2018). In addition, phonological neighbourhood density has a particular effect on voiceless stop VOT, in that VOT is enhanced for lexical items with a stop-initial competitor and overall high phonological neighbourhood density (Baese-Berk and Goldrick 2009; Kirov and Wilson 2012; Fox et al. 2015; Nelson and Wedel 2017). Similar VOT enhancement is observed in misheard speech in which the competitor word differs from the target word in word-initial stop voicing only (Schertz 2013). Knowledge of the lexicon can also influence perception of VOT (Ganong 1980), in which listeners are more likely to perceive a real word than a non-word, especially at ambiguous VOT values. VOT can also influence lexical access: listeners are slower to access words with a laryngeal competitor (*pat vs. bat*) than those without (*king* vs. *ging*), and the exact VOT can affect the access time at least in early stages of perception (Andruski et al. 1994).

Finally, syllable stress and the number of syllables within a word have minor effects on voiceless stop VOT. VOT is longer preceding stressed vowels, but this difference is diminished in running speech (Lisker and Abramson 1967; Klatt 1975; Keating 1984). Word-initial voiceless stop VOT is also longer in monosyllabic words than polysyllabic words (Klatt 1975; Flege et al. 1988).

2.5. Phonological alternations

Stops are subject to a wide range of phonological processes in which articulation is affected in quantal ways such that we might model them as segmental deletions, additions, or substitutions rather than adjustments to the gestural phases of stop production. These processes vary between and within languages, and are often linked to social differences between speaker groups. Documenting these processes comprehensively is beyond the scope of this work, and we cite just a few examples here.

In English, for example, /t/ can be realised as a range of variants depending on phonological context and accent. These variants include [ɾ d ɹ ʔ], full deletion (*must go* > [mʌs ɡəʊ]), and assimilation to following bilabials or velars (*hot coffee* > [hɒk͡ kɒfi]). The linguistic constraints on variant patterns can be very complex. The glottal variant of /t/ in British English varies in frequency according to at least the following: word and syllable position, following segment, lexical status (more glottals are used in grammatical words than lexical words), lexical frequency (more glottals in more frequent words, as is typical for lenition processes), and whether the word is repeated in discourse (Hay and Foulkes 2016, Foulkes and Hughes in press). It is also subject to considerable social and regional variation (see 3.4 below). Another widely studied variable in English is /t, d/ deletion in word-final consonant clusters (Tagliamonte and Temple 2005, Raymond, Brown and Healy 2016, Baranowski and Turton 2020, MacKenzie and Tamminga 2021). Most studies find that deletion is most common in monomorphemes (*mist*) and least common in regular past tense forms (*missed*), with irregular past tense forms (*kept*) accepting deletion at an intermediate rate. The nature of the following segment is also crucial (more deletion before consonants, less or no deletion before vowels or pause), as are various higher-level factors relating to morphological structure. Evidence from articulatory studies, however, suggests that complete deletion might be an auditory/acoustic rather than physical effect, with evidence that some articulatory gesture tends to remain even in cases where there is no audible stop (Purse 2019). Deletion processes occur in many languages, including Suva Rotuman (deletion of /ʔ/ in initial position; Fimone 2020), Capanahua (deletion of /ʔ/; Elias-Ulloa 2004), Persian (coda obstruent cluster simplification; Falahati 2013), and Catalan (word-final stop deletion; Mascaró 1989).

The opposite of deletion, epenthesis, is also observed in many languages, for example in nasal+fricative sequences in Dutch and English (e.g. *prince* realised as [pɹɪnts], *Chomsky* as [tʃɒmpski], or *strengths* as [stɹɛŋkθs]). Epenthesis is most common in word-final position (Yoo and Blankenship 2004). Epenthetic stops are shorter than phonological stops (Fourakis and Port 1986). As such they can be analysed as epiphenomenal, generated by mistiming of the articulatory gestures for the nasal and fricative, rather than segmental insertions. In perception experiments. Warner and Weber (2001, 2002) found that listeners were slower to identify epenthetic stops than lexical stops, especially in medial position.

Segmental alternations are found in most if not all languages. Assimilation, which is generally anticipatory, can affect place, manner and/or voicing (Zsiga 2011). As some examples, assimilation of place and manner is reported, for instance, in Catalan (Recasens and Mira 2015), voicing assimilation in Russian (Padgett 2002), and place assimilation in Korean (Kochetov and Pouplier 2008). Voiced stops (and other obstruents) may be subject to devoicing in certain contexts, including word-final position, adjacency to voiceless consonants, and preceding pause (for a cross-linguistic review see Keating, Linker and Huffman 1983). Such patterns are observed, for example, in German (Wiese 2000), Wolof (Ngom 2003) and Turkish (Becker et al. 2011). Devoicing in word initial position occurs in Japanese (Gao and Arai 2019).

3. Regional and social variation in stops

In this section we focus on within-language patterns that can be considered ‘sociophonetic’ in that they differentiate regional or social subgroups within a population. In line with section 2, we organise the discussion around the phonetic phases of stop production, with separate sections on VOT and phonological processes.

3.1. Closing phase

Pre-aspiration patterns with speaker sex in several languages, with females producing more and longer preaspiration than males. This is the case, for example, in Nordic languages (Helgason 2002, Stölten and Engstrand 2002, Helgason and Ringen 2008), Welsh (Morris and Hejná 2020) and Newcastle English (Docherty and Foulkes 1999). Younger speakers have shorter and less intense preaspiration than older speakers of Gaelic (Nance and Stuart-Smith 2013). Prenasalisation of Greek stops appears to vary by region, younger speakers produce it less than their elders, and it is less common in more casual styles (Arvaniti 2007).

Incomplete closure - one form of lenition - typifies accents of English and Italian, among others. In Liverpool English /k/ is stereotypically realised as [x], especially in non-initial contexts, although all stops are affected by lenition, subject to constraints of word and syllable position (Clark and Watson 2016). The lenition pattern is spreading outwards from Liverpool to surrounding areas, but with simplification in the linguistic constraints operating on some variants. Irish accents often display a fricative realisation of /t/ in coda position, often referred to as ‘slit-T’ (O’Dwyer 2019). Detailed analysis shows that slit-T is usually voiceless and apico-alveolar, but a wider range of variants may be used including voiced fricatives, a laminal [s], and flaps (Skarnitzl and Rálišová 2022). In Italian, lenition of /p t k/ to frictives is characteristic of Rome, and central and southern varieties (Vietti 2019). The process is expanding, with signs that it holds a degree of prestige and signals a macro-regional identity. Lenited realisations of stops are also stereotypical of Tuscan accents (a process known as *gorgia toscana*), especially the high status accent of Florence. Stops lenite in increasing severity to fricatives, approximants, [h] and full deletion, depending on speech rate and stylistic register. The Australian language Murrinh Patha displays an unusual lenition pattern affecting /p/ and /k/ in stressed onsets, and young men lenite more than other speaker groups (Mansfield 2015).

Variation in place of articulation is well known as an accent feature in English. Dental variants of /t, d/ are stereotypical of some Irish and Caribbean accents (Wells 1982). Similar but socially subtler variation is found in South Africa (Mesthrie 2012). Dental variants are common in Cape Town but not Durban, while all coronals are variably dental for broader Cape Town Coloured and Indian speakers. Retracted or retroflex variants are common for some speakers of British Asian accents (Kirkham 2011). In Multicultural London English, /k/ is often backed towards [q] (Fox and Torgersen 2018).

3.2. Hold phase

We can find relatively little evidence of social or regional variation in the hold phase of stops. A number of studies report small differences in hold duration related to age, but this appears biologically rather than socially driven (e.g. Benjamin 1997).

Some regional variation is found in geminate consonants, including stops. Degemination occurs in northern Italian accents, while central and southern accents observe a pattern of consonant lengthening via external sandhi in word-initial positions when preceded by a final stressed vowel or certain specific morphemes (Vietti 2019). For example, *tre case* 'three houses' [ˌtreˈkːaːse]. This process is called *raddoppiamento sintattico* ('syntactic doubling'). Vietti (2019) comments that this process has only rarely been subject to sociolinguistic investigation. However, in Turin it appears that *raddoppiamento sintattico* is spreading among young immigrants from the Maghreb, apparently via contact with southern speakers (Boario 2009, cited by Vietti 2019). Vietti suggests it may therefore eventually become a sociolinguistic marker. Degemination is stigmatised in Veneto (Trumper and Maddalon 1990, cited by Vietti 2019). Gemination is reported as a feature of Cypriot Greek, but not standard Greek (Alexander 2016).

Bundgaard-Nielsen and O'Shannessy (2021) report an acoustic study of stops in Light Warlpiri (a mixed language spoken in the Northern Territory of Australia). Speakers show an interesting phonological development arising from contact between the indigenous language Warlpiri and both English and Kriol (i.e. a creole largely lexified from English). While Warlpiri has a single stop series, consisting of voiceless stops at five places of articulation, the language shows amalgamation of the phonological categories of both Warlpiri and English/Kriol. VOT is emerging as a contrastive feature in initial positions, while in medial position, constriction duration (i.e. the length of the hold) is the main cue for contrast. Short constriction duration is used for English and Kriol words with a phonologically voiced stop, whereas longer duration is used to signal voiceless stops.

3.3. Release phase

A small number of studies have considered regional or social variation in intensity or spectral frequency characteristics. Kirkham (2011) found that British Asian English /t/ was marked by a greater relative burst intensity than corresponding /d/, whereas British White English speakers demonstrated the opposite pattern. Measures of spectral shape, such as COG, are commonly used for studying fricative variation, but only occasionally for stops. Female speakers tend to have a higher COG for stops relative to male speakers in English (Sundara 2005; Chodroff and Wilson 2014). British Asian English speakers produce /t/ with a higher COG, skewness, and kurtosis, as well as a lower spectral SD, relative to British White English speakers. In Derby, UK, impressionistic comments suggest older speakers produce lower frequency releases of /t/ and /d/, which might reflect a more retracted articulation relative to younger speakers (Docherty and Foulkes 1999: 51). We return to this issue in our case study in section 4.

The acoustic properties of release bursts are furthermore reported to separate (female) Modern Standard Greek and Cypriot Greek speakers (Themistocleous 2016). Aspiration of voiceless stops appears in Yucatan Spanish (Michnowicz and Carpenter 2013) and some non-standard accents of southern Italian (Vietti 2019). Nodari, Celata and Nagy (2019) explore change in long lag VOT in Calabrian heritage language speakers in Toronto.

Contact between languages has been cited to explain developments over time in VOT patterns. For example, contact with standard Spanish has led to loss of aspiration for Patagonian Welsh (Sleeper 2018), and contact with English has led to longer aspiration in Māori (Maclagan et al. 2009). Wang (2019) explores contact between Jin and Mandarin in Hohhot, a city in Inner Mongolia. The city is split into an old town and new town, the former being Jin-speaking, while the latter is populated by incomers from all over China who mainly use Mandarin as a *lingua franca*. Jin speakers typically realise the fricative element of aspirated stops /p^h t^h k^h/ with a back fricative, [^{x}] (and /h/ itself as [x]). The new town speakers are developing a new contact variety, in which they converge on the [^{x}] variant. Sex and age effects are also found.

Stop release rate in coda position has also been investigated. Davidson (2011) reports a strong ethnic difference in the rate of stop releases in word-medial, word-final, and phrase-final positions: white speakers had a higher rate of stop releases than black speakers of American English. Several studies on English also consider the release patterns of final /t/ as cues to structure conversation. For example, fully released utterance-final [t] acts as a turn-handover cue for speakers in Tyneside (Docherty et al. 1997), while ejectives may play a similar role (Ogden 2009). Variants of /t/ can also be harnessed to signal aspects of identity or persona. Podesva et al. (2015) suggest that released [t] is used as a marker of an 'articulate persona' among US politicians, while the same feature is used to construct 'nerd' identity (Bucholtz 1996) and used as a symbolic marker in constructing Jewish identity (Benor 2001, Levon 2006) or to 'sound gay' (Podesva et al. 2001). Palatalized /t/, in combination with fronted /s/, increases percepts of gayness in Copenhagen Danish (Pharao and Maegaard 2017). In French, affricated /t/ is a stigmatised feature of *cité* (referring to low status housing) or *beur* (North African heritage) speech, but holds covert prestige due to the prevalence of these urban varieties in hip hop (Jamin 2004, Gadet and Hambye 2014). The feature appears to have spread into Moroccan Arabic hip hop and rap styles (Schwartz 2018).

3.4. VOT and laryngeal timing

As mentioned in section 2.4, VOT varies considerably across languages. Even within a language, broadly construed, regional differences are reported in several languages including Danish (Puggaard 2021), French (Caramazza and Yeni-Komshian 1974), German (Kleber 2018), Italian (Villafaña Dalcher 2008, Vietti 2019), Korean (Cho et al. 2002), and Japanese (Utsugi et al. 2013; though the evidence in this case is inconsistent). In English, speakers of some regional accents use shorter VOT for /p t k/ than is found in standard accents. They include Scottish varieties (Scobbie 2006, Docherty et al. 2011, Stuart-Smith et al. 2015), some accents in northern England, notably in Lancashire (Wells 1982: 74, 370), and South African varieties (Bowerman 2008, Finn 2008, Mesthrie 2008a, van Rooy 2008). Regional varieties of English in several parts of Asia also manifest deaspiration, presumably through the influence of other languages spoken in the respective location (Gargesh 2008, Mahboob and Ahmar 2008). Prevoicing for /b d ɡ/ is reported for some speakers of many languages in which the standard variety uses short lag. This includes English (e.g. Lisker and Abramson 1964, Docherty 1992). Prevoicing varies regionally across the US, and is used more by African Americans than Caucasians (Herd 2020).

Variation in VOT also correlates with social factors in several studies. It is often found that women produce longer aspiration than men (e.g. English: Swartz 1992; Ryalls et al. 1997; Whiteside et al. 2004; Mandarin: Ma et al. 2018; Mandarin and Hakka: Peng et al. 2014). By contrast, men produce longer unaspirated VOT (e.g. Li 2013; Peng et al. 2014, Ma et al.

2018) and use more pre-voicing in phonologically voiced stops (e.g. Ryalls et al. 1997 for US English, Helgason and Ringen 2008 for Swedish). Children demonstrate sex differences in VOT production along the same lines as adults (Whiteside and Marshall 2001; Whiteside et al. 2004).

The exact factors which contribute to these sex differences are yet to be identified, although aerodynamic, physiological and anatomical factors have all been presented as suggestions (Koenig 2001; Swartz 1992). However, the sex-based pattern is not universal. Morris et al. (2008) found no difference between American men and women when controlling speech rate; Li (2013) also found that the Mandarin gender difference in aspiration could be accounted for by differences in speaking rate. Moreover, it is men who produce longer VOT in Florentine Italian (Piccardi 2017), Hungarian (Gósy and Ringen 2009), Korean (Oh 2011), Mandarin (Li 2013) and Serbian (Sokolovic-Perovic 2012). Thus we must infer that sociolinguistic factors must also play a role in shaping VOT production.

Differences in VOT act as indicators of ethnicity among Lebanese Australians, who use more prevoicing and slightly shorter aspiration than comparator speakers (Clothier and Loakes 2018), while VOT patterns among Polish speakers of English appear linked to issues of 'nationalist' versus 'cosmopolitan' identity (Kozminska 2015). The latter group display longer VOT for /p t k/, reaching native English norms, while the 'nationalist' group retain short VOT influenced by underlying Polish norms. Ryalls et al. (1997) and Herd (2019) also found more prevoicing for young African Americans than young Caucasians.

There is conflicting information as to whether VOT increases or decreases as speakers age, although speech rate generally slows for elderly speakers and thus VOT lengthens as a consequence (Stuart-Smith et al. 2015). However, some studies report no difference by age in VOT means (Neiman et al. 1983; Petrosino et al. 1993), whereas others report that older speakers have reliably shorter VOTs than younger speakers (e.g. Benjamin 1982; Morris and Brown 1994; Ryalls et al. 2004; Torre and Barlow 2009; Docherty et al. 2011; Kleinschmidt et al. 2018; Ma et al. 2018). Thus the age differences cannot be solely attributed to speech rate differences. In addition, Torre and Barlow (2009) outline the importance of considering gender-specific age differences in evaluating VOT and speech. Age-related changes to the body differ by sex: changes to laryngeal structures tend to be greater in men, whereas changes to the respiratory system tend to be greater in women (Gorham-Rowan and Laures-Gore, 2006). Increased VOT variability has also been found in both children and elderly speakers (Petrosino et al. 1993). VOT variability decreases with age up until adulthood (Zlatin and Koenigsknecht 1976; Macken and Barton 1980; Clumeck et al. 1981; Koenig, 2001; Yang 2018). Effects of parental input on VOT are reported for example by Scobbie (2006), who found that Scottish teenagers with Shetland Island parents tended to show at least some use of the Shetland pattern, in which the voicing contrast is signalled by pre-voicing in /b d g/ versus short lag VOT in /p t k/.

Further confounds are introduced by speech style: VOT tends to be longer in careful speech such as word-list readings, and shorter in spontaneous speech (Lisker and Abramson 1967, Baran et al. 1977). VOT is also reported to be longer for /p t k/ in some (but not all) studies of infant-directed speech (Fish et al. 2017); however, these studies have mostly been restricted to US English. Data for English /t/ from Fish et al. (2017), for example, reveals a mean of 130 ms in Infant-directed speech compared to 101 ms in adult-directed speech when analysed from words in context, i.e. VOT is around 30% longer in speech to infants.

A number of studies indicate that VOT can be subject to change over time. Traditional RP had rather shorter aspiration than its modern descendant (Hickey 2017). The standard German VOT distinction has been lost in some dialects, but the merger appears to be in the process of reversal in Bavaria and Saxony, likely due to influence from the standard (Kleber 2018). In Korean, VOT for aspirated and lenis stops has increased over time, and f0 in the following vowel has emerged as an important cue to contrast (Lee et al. 2020; see also section 2.4). Social and regional variation is also reported for devoicing patterns in French (Temple 1999) and Setswana (Duran et al. 2017).

Other aspects of timing show regional and social variation. Tyneside English, for example, has a very distinctive set of local variants for intervocalic /p t k/, which are generally signalled by partial or full voicing during closure and a secondary glottal constriction and release prior to or following the release (Docherty and Foulkes 1999, 2005). The supralaryngeal constriction in these variants is often lenited. Moreover, the timing of the secondary glottal articulation varies between speakers, with older men in particular tending to produce the glottal earlier in the sequence. This in turn leads to a greater proportion of clear oral release, which for younger speakers is often masked by the secondary glottal action. Interpreting change is compromised by the difficulty in controlling both speech rate and speaker age, however.

The degree of VOT imitation and convergence can also be modulated by social factors. Sanchez et al. (2010) found that speakers (all female US students) modulated VOT to be more similar to their interlocutor in a speech shadowing task. The degree to which speakers converge to an interlocutor can also depend on the speaker's attitude towards the interlocutor (Abrego-Collier et al. 2011; Yu et al. 2013). Adolescents are less likely to converge on VOT than adults (Schertz and Johnson 2022). Moreover, there is considerable individual variability in the degree of imitated convergence (Wade et al. 2021).

Several social factors have been implicated in the production and perception of the Korean three-way stop contrast. Younger Seoul listeners are more likely to rely on f0 for the lenis–aspirated contrast than older Seoul listeners (Kang 2010), and Kyungsang listeners are more likely to rely on VOT for that contrast than Seoul listeners (Lee et al. 2013); these perceptual patterns are largely mirrored in production (Kang and Guion 2008; Lee and Jongman 2012). Kong, Holliday and Lee (2022) identified more flexibility in the relative weighting of VOT and f0 for the Korean three-way stop contrast among university speakers than younger students, who were more likely to perceive the contrast in line with their local production patterns.

3.5. Phonological alternations

The types of phonological alternations we discussed in section 2.5 often display sociolinguistic patterning. For instance, one of the most salient sociolinguistic markers in Arabic is (q), referring to the variants [k g q ʔ] used for standard /q/ (Watson 2002: 17; Khattab and Foulkes, this volume). Variation is complex, related not only to region but to sex, age, ethnicity and a rural-urban distinction. Regional variants typically include [g] in Jordan, Upper Egypt and Saudi Arabia, [ʔ] in Levantine varieties and Lower Egypt, [k] in the rural West Bank, while [q] dominates in Maghrebi and rural Syrian dialects, and amongst the Druze community in Lebanon. In Palestine, [k g ʔ] are subject to ongoing change linked to age, sex and dialectal background (Cotter 2014). Usage patterns are further complicated by the diglossic situation, whereby [q] is used in most dialects for learned words and when speakers switch to Modern Standard Arabic.

In English, variants of /t/ have probably been the subject of more studies than any other consonant. Glottaling of /t/, particularly in intervocalic position, is generally associated with non-standard accents in the UK, as well as less formal styles and speakers of lower social status. Men generally use more glottals than women, as is typical of sociolinguistic changes from below (e.g. Straw and Patrick 2007, Schleef 2013; but see also Mees and Collins 1999). Glottal variants are spreading in RP (Fabricius 2002), reaching even as far as the speech of younger royals (Shaw and Foulkes 2015). Several recent studies have also assessed the emergence of the glottal variant of /t/ in US English, where the change appears to be led by young women (Ellingson Eddington and Brown 2021). /t/ variants can generate strong social evaluations by listeners. The glottal variant is generally stigmatised in the UK, although the stigma appears to be waning fast (Fabricius 2002). In the US both glottals and full releases received negative evaluation in the study by Ellingson Eddington and Brown (2021): speakers who used glottal stops were viewed as less educated and less friendly, while full oral releases were judged more ‘rustic’ and less educated. In Tyneside there is evidence that children are exposed to sociolinguistic variation differently, depending on their age and sex. Girls typically hear more standard variants of /p t k/ than boys do (Foulkes et al. 2005).

In North American and Australasian Englishes, intervocalic /t/ is generally realised as a tap, [ɾ]. The change towards this variant from [t] is tracked in detail for New Zealand English by Hay and Foulkes (2016), who show effects of age, sex, social class, word frequency, ‘word age’ (i.e. words used more by younger speakers contain more of the innovative variant) and topic of conversation. In Australia, /t/ variants differ across Aboriginal and ‘mainstream’ communities (Loakes et al. 2018).

Epenthesis of /t/ in English shows variation both in respect of individual speaker (Yoo and Blankenship 2003) and also region: although common in many varieties, it appears not to occur in South African English (Fourakis and Port 1986). Deletion of /t, d/ in obstruents is regionally and socially variable, occurring in pre-vocalic and pre-pausal positions in African American English as well as the more common pre-consonantal context observed in standard accents (e.g. Labov 1972). /t, d/ deletion has also been studied in Jamaican creole (Patrick 2008) and Hispanic varieties in the US (e.g. Bayley 1994, Santa Ana 1992). As for assimilation processes, the anticipatory place of articulation assimilation found in standard English appears not to occur in Durham, northern England (Kerswill and Wright 1990), which by contrast permits regressive voicing assimilation. A similar process is found in West Yorkshire, thus Hyde Park may be realised as [haɪt̚ paːk] (or [haɪʔ paːk] via glottaling), although this feature seems to be receding (Whisker-Taylor and Clark 2019).

Deletion of initial /ʔ/ in Rotuman shows social variation and change in progress (Fimone 2020). The deletion process is used less by younger speakers and older women, suggesting increased influence of the standard form as well as social identity factors at work. Noglo (2009) investigates variation and change in Ewe, as spoken in Lome, Togo. Among the variables reported is fortition of /ɸ/ to [p], the latter being an urban rather than rural marker and used more by women than men. Regional variation in stop realisation is also reported for Turkish (Yağlı 2018).

3.6. Summary

In the previous two sections we have provided a summary of sources of variation and change affecting stops. As shown in section 2, many phonetic studies have shown systematic variation relating to age, speaker sex, and linguistic-phonetic factors such as structural position and speaking rate. Explanations for differences in stop realisation have therefore often been offered based on differences in biology or physics. However, as we have seen in section 3, social and regional factors are also often apparent. Given the physical complexity of stops, there is huge potential for further sociophonetic analyses of stops in all languages. Auditory and acoustic analyses are often sufficient for such work, but it is notable that newer techniques such as ultrasound are also being applied to non-standard varieties and diverse speaker groups.

In the final section we turn to a case study in which we investigate systematic social variation of sex and age on the phonetic realisation of stop consonants in spontaneous Derby English. As with many of the studies referenced above, we investigate the extent to which sex and age influence aspects of stop realisation. We additionally investigate the interface between social variables and phonological features in the process of phonetic realisation.

4. Case study

A key component of sociophonetic analysis concerns characterising the relationship between social/regional and phonetic/phonological variables. Phonetic/phonological variables are traditionally characterised either as segmental units, such as the use of the variant [ʔ] for English /t/, or as a specific acoustic or articulatory dimension of a single speech sound, such as the VOT of [t^h] or the COG of [s]. Previous sociophonetic studies have therefore generally focused on the relationship between one or more social variables and the phonetic variants of the unit of analysis, or have grouped a natural class of sounds together under the assumption that they *should* pattern together (e.g. VOT of the voiceless stop series, [p^h t^h k^h]). This assumption of natural-class patterning, however, implicitly assumes that the relevant social variable targets a phonological feature shared by all members of the class.

In this case study, we investigate the relationship between traditional social and phonetic variables, and also take a closer look at the exact relationship between them. Is it the case that social variables affect single speech sounds, or rather, do social variables affect the phonetic realisation of a natural class of speech sounds? Social variables can certainly target a single speech sound. For example, in English, /t/ is known to take on a variety of phonetic realisations and undergo several phonological alternations that correlate with social factors; this modulation happens largely to the exclusion of related voiceless stops, /p/ and /k/, or the related coronal stop, /d/. In some phonological models, this exceptionality is attributed to a notion of markedness or underspecification, and specifically, that /t/ could be underspecified for either place or voicing features, making it ripe for malleability in its realisation (e.g. Avery and Rice 1989; Paradis and Prunet 1991; Backley 1993; Lombardi 2002). Just as plausibly, however, a social variable could instead target a shared phonological feature, thus influencing the phonetic realisation of multiple speech sounds simultaneously. Phonetic realisation here refers to the mapping between a phonological surface segment to an idealised plan of phonetic targets (e.g. Keating 1990; Cohn 1993; Fruehwald 2017; Volenec and Reiss 2017). For example, Fruehwald (2017) suggested that the fronting of Philadelphia English vowels over time may actually target the shared [+back] feature: it is not that individual vowels front, but rather that the class of phonologically back vowels shifts in parallel to one another. Synchronically, this might also play out as a *uniform* phonetic realisation of a distinctive feature value across all segments specified with that feature value (e.g. Chodroff and Wilson 2017). A pressure for uniformity may arise from a more general linguistic constraint of target uniformity: this refers to near-identity in the phonetic targets corresponding to a shared distinctive feature across multiple speech sounds. In support of this, Chodroff and Wilson (2018) identified overall phonetic variation in the overall realisation of stop consonants across American English speakers; however, the expression of a shared laryngeal feature, as measured by VOT, was relatively uniform within a speaker, as was the expression of a shared place of articulation feature, as measured by spectral centre of gravity (COG).

Note that this question also has relevance for usage-based models such as exemplar theory, which have been applied in several studies of variation and change. While usage-based models assume that abstract units emerge via analysis of experienced and memorised events, it remains unclear what form(s) of abstraction take place and what phonological primes and processes emerge (see e.g. Docherty and Foulkes 2014). It thus also remains unclear how social-indexical values emerge: are social variables attached to specific exemplars (or ranges of exemplars), and/or to more abstract units? (For further discussion see e.g. Labov 2006, Pierrehumbert 2006, Docherty and Foulkes 2014, Foulkes and Hay 2015.)

The interface—and interplay—between social/regional variables and phonetic realisation, alongside linguistic constraints and phonetic realisation is a fruitful area for the development of sociophonetics. Phonetic realisation is subject to multiple pressures: on the one hand, speakers may express aspects of their social identity through this process, but on the other hand, phonetic realisation may be highly constrained by more general linguistic factors, such as target uniformity (see also Guy and Hinskens 2016 on social and linguistic coherence). Target uniformity ensures the phonetic realisation of a feature; the corollary is that if the realisation of the feature is consistent across segments, then the realisation of the contrast is preserved across pairs of segments. If the laryngeal targets corresponding to [p^h t^h k^h] are the same, as are those corresponding to [b d g], then the contrast between [p^h b], [t^h d] and [k^h g] will also manifested by consistent phonetic cues. Because of this, we hypothesise that the expression of social information should be subordinate to preserving the structure of phonological contrasts, and that social factors should have minimal to no interference with target uniformity.

In the present study, we examine the influence of social factors and linguistic constraints on the phonetic realisation of stop place of articulation and voicing in a spontaneous speech corpus of English from Derby, UK. The phonetic realisation of stop place of articulation will be approximated by its spectral shape, a primary acoustic correlate, and specifically the COG of the initial release burst. COG is the energy-weighted mean frequency of the spectrum. While stop COG has been implicated in stop voicing contrasts, these differences are considerably smaller than the differences across place of articulation (see section 2.3). The phonetic realisation of stop voicing will be approximated by positive VOT (see section 2.4). Important to note is that the acoustic measurements are not the same as the phonetic targets, but the phonetic targets can be approximated via the acoustic instantiation. For instance, VOT largely reflects the timing between the glottal spreading gesture and the oral release. These gestures and their timing relationship could reasonably serve as the phonetic targets corresponding to the phonological voicing feature. The analysis examines all six stop consonants, [p^h t^h k^h b d g], in word-initial position from 35 native Derby English speakers, balanced by sex and age.

Given the above hypothesis, the largest influences of social factors should be on the entire class of stops or on the primary phonetic correlate to a phonological contrast. Changes along these dimensions would still preserve the structure of phonological contrasts in phonetic realisation (e.g. VOT differences between voiced and voiceless stops). Social factors should have minimal influence on phonetic targets relating to a non-primary distinctive feature (e.g. laryngeal phonetic targets for phonological place features), as this would violate target uniformity.

VOT serves as the primary correlate to the voicing contrast, and only a secondary correlate of the phonological place of articulation contrast. We would expect social factors to have larger influences on the VOT contrast between voiced and voiceless stops than among stops differing in place. Similarly, as COG primarily reflects place of articulation and secondarily reflects voicing, we expect social factors to have a larger influence on the realisation of place than of voicing with respect to COG. In addition, we do not expect there to be many three-way interactions between place, voice, and social factors, in which individual segments are targeted by social factors.

4.1. Methods

4.1.1. Participants

35 speakers from Derby were recorded in fieldwork conducted in 1995. The speakers were reasonably balanced across demographic factors of sex (female or male), age (younger or older), as well as professional status (broadly 'working class' and 'lower middle class'). The present study focuses on sex and age. (Professional status was more difficult to balance in a categorical manner, and was not considered in the present study. Indeed, many social variables may be better captured through continuous measures based instead on degree of self-reported affiliation or factual degrees like exact age.) Younger speakers were born between the years 1967 to 1981 and were 14–27 years of age at the time of recording. Older speakers were born between 1925 and 1957 and were 38–69 years of age at the time of recording. Factors of sex and age were counterbalanced across 32 of the 35 speakers (8 female–younger, 8 male–younger, 8 female–older, 8 male–older). Three elderly speakers (2 female, 1 male) were also recorded; these speakers were born between 1913 and 1916 and were 79–82 years of age. They are included here in the respective older speaker groups.

4.1.2. Procedure

Each conversation took place in self-selected dyads among the speakers (in the case of the elderly speaker group, a triad). Within groups, age and professional status were matched, but the pairing of sex varied. The total speech duration per speaker ranged from 10 to 37 minutes with an average of 24 minutes of speech per speaker. Recordings were made with 46-minute Digital Audio tapes and digitised with a sampling rate of 22,050 Hz (Docherty and Foulkes 1999).

4.1.3. Processing

The corpus includes audio, various alignments and critically a word- and phone-level alignment, along with meta-data for the participants; annotations were completed using LaBB-CAT (Fromont and Hay 2012). AutoVOT was used to identify the stop release and following vowel onset boundaries using automatic methods (Keshet, Sonderegger and Knowles 2014). The present analysis focused on word-initial, prevocalic stop consonants, [p^h t^h k^h b d g]. Following Chodroff and Wilson (2017), the phone-level alignments of voiceless stop consonants were extended 31 ms beyond each pre-aligned boundary, and for voiced stops, 11 ms (see Figure 1). The minimum permissible VOT value was set to 15 ms for voiceless stops and 4 ms for voiced stops. Stops were removed if they occurred in an unfinished word or in the word “to”, due to its common reduction. A total of 16,032 stops met the initial criteria for analysis. Of these, 1,610 stops (~10% of the corpus) were audited and corrected based on random sampling or a suspicious VOT value (e.g. excessively long or the minimum threshold). Audited stops without a release burst were removed (n = 232), leaving 15,797 stops available for analysis (see Table 1 for a summary of counts per stop and speaker).

a) b)

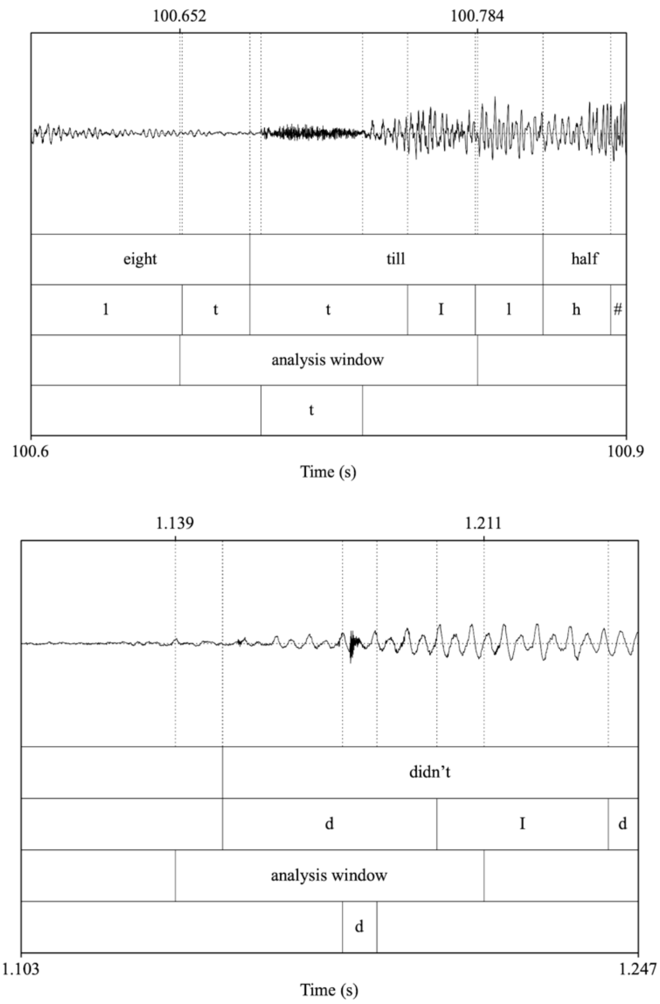

Figure 1. Analysis windows on Tier 3 which were submitted to AutoVOT for a) a word-initial voiceless stop and b) a word-initial voiced stop. For voiceless stops, the window started 31 ms before the force-aligned stop onset (Tier 2) and ended 31 ms after the force-aligned stop offset. For voiced stops, the analysis window started 11 ms before the force-aligned onset and 11 ms after the force-aligned offset. AutoVOT searched for the stop release and onset of periodicity within this window.

Stop	Total	Median	Range
[p^h]	1247	37	4–60
[t^h]	1597	43	8–127
[k^h]	2302	65	11–139
[b]	4343	126	16–272
[d]	3319	92	21–203
[g]	2989	75	18–187

Table 1. Total number of tokens available for each stop category, along with the summary counts per speaker.

4.1.4. Measurements

Positive VOT and COG from the initial release burst were extracted from word-initial, prevocalic stop consonants. Positive VOT was measured as the duration between the automatic or manually-corrected boundaries described above. COG is the energy-weighted mean frequency of the spectrum, and was taken from a time-averaged spectrum over the initial stop burst. The smoothed spectrum was the average of seven 64-point FFT spectra that were extracted from 3 ms Hamming windows with a 1 ms window shift. The first window was centred 1 ms into the initial release (Hanson and Stevens 2003; Chodroff and Wilson 2014). If the stop was shorter than the amount of time required for the full seven windows, then the number of windows was reduced to accommodate the duration. Prior to extraction, the audio was high-pass filtered at 200 Hz and pre-emphasized above 1000 Hz (see also Forrest et al. 1988; Sundara 2005). Measurements were extracted from the manually corrected stop consonant when available.

4.1.5. Analysis

Variation in VOT and COG from the initial release burst was analysed using Bayesian mixed-effects linear regressions (brms: Bürkner 2018). Each linear regression had the same set of predictors; all categorical predictors were sum-coded. VOT and COG were predicted from fixed effects of place of articulation (*coronal*: coronal +1, dorsal 0, labial -1; *dorsal*: coronal 0 dorsal +1 labial -1), voicing (voiceless +1, voiced -1), following vowel height (high +1, non-high -1), following vowel duration (ms), sex (female +1, male -1), age (older +1, younger -1), and the full interactions between the phonological and social predictors (place, voicing, sex and age). The model for VOT included an additional interaction between voicing and following vowel duration, as speaking rate is known to influence voiceless stop VOT more than voiced stop VOT (Kessinger and Blumstein 1997). The random effect structure included a by-participant intercept and slopes for place and voicing.

Prior distributions over fixed effects were estimated where possible using Mixer-6 VOT data from 180 speakers of American English released with Chodroff and Wilson (2018) (https://osf.io/jt5mc/). All other priors were only weakly informative. The exact model specifications, raw data, code, and model results can be found in the OSF repository (https://osf.io/6a5jq/).

Each beta coefficient can be interpreted as the deviation from the estimated mean VOT or COG in milliseconds and hertz respectively. This is because categorical predictors were sum-coded and continuous predictors were centred. Doubling the beta estimate yields the estimated contrast between binary predictors. Reliability of an effect was determined using the 95% credible interval (CI) over the marginal posterior distribution for the coefficient estimate. If the 95% CI excluded 0, then we could conclude that the direction of the estimate was reliable in the estimated direction with probability < 0.025.

4.2. Results

4.2.1. VOT

Variation in VOT by sex, age, and stop category is shown in Figure 2 and Table 2. The model of VOT variation revealed credible effects of several linguistic and social factors (see Figure 3). As expected, voiceless stops had longer VOTs than voiced stops. The model generates the following coefficient: voice: 16.95, 95% CI [15.90, 18.00]. In line with the explanation given at the end of 4.1, this means that voiceless stops were approximately 17 ms longer than the average VOT, and voiced stops were approximately 17 ms shorter than the average VOT, yielding an overall difference of 34 ms. Coronal and dorsal stops had overall longer VOTs relative to the average (coronal: 1.45, 95% CI [0.63, 2.27]; dorsal: 2.17, 95% CI [1.35, 2.94]), but this was reliably modulated by voicing. Compared to estimates from the main effects of place and voicing alone, the voicing contrast was reliably larger for coronal stops and reliably smaller for dorsal stops (coronal × voice: 2.55, 95% CI [1.83, 3.26]; dorsal × voice: -1.05, 95% CI [-1.87, -0.21]). Longer VOTs were also observed prior to longer vowel durations, particularly for voiceless stops (vdur: 30.69, 95% CI [27.74, 33.70]; voicing × vdur: 5.49, 95% CI [2.54, 8.56]). The effect of vowel height was not reliable: VOT before high vowels was not reliably different from VOT before non-high vowels (vowel: -0.02, 95% CI [-0.48, 0.43]).

Among social factors, sex and age reliably structured VOT variation: female speakers had longer VOTs than male speakers, and younger speakers had longer VOTs than older speakers (sex: 3.11, 95% CI [1.69, 4.51]; age: -2.81, 95% CI [-4.25, -1.33]). Moreover, the voicing contrast was credibly larger for female than male speakers, larger for younger than older speakers, and the contrast further reduced specifically for older men (voicing × sex: 2.27, 95% CI [1.23, 3.31]; voicing × age: -2.56, 95% CI [-3.64, -1.48]; voicing × sex × age: 1.35, 95% CI [0.31, 2.38]). Finally, female speakers had longer VOTs for coronal stops than male speakers relative to the expected coronal stop VOT (coronal × sex: 0.82, 95% CI [0.01, 1.62]). Moreover, female speakers had an increased contrast between [t^h] and [d] VOT relative to male speakers (coronal × voicing × sex: 0.84, 95% CI [0.13, 1.53]). No other interactions were reliable in their direction.

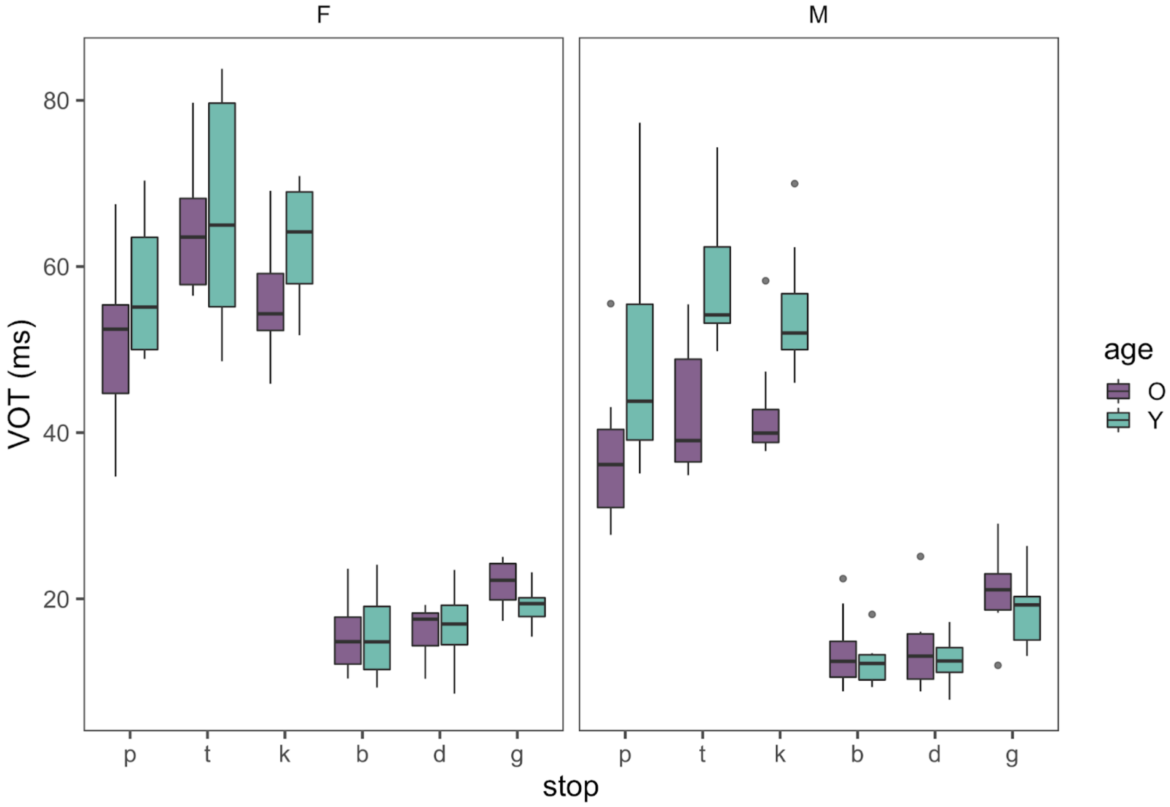

Figure 2. Variation in by-speaker VOT means (ms) for [pʰ tʰ kʰ b d g] across sex and age.

VOT	**[pʰ]**	**[tʰ]**	**[kʰ]**	**[b]**	**[d]**	**[g]**
Younger female	57 (8)	67 (14)	63 (7)	16 (5)	17 (5)	19 (3)
Younger male	49 (14)	58 (8)	54 (8)	12 (3)	13 (3)	19 (4)
Older female	51 (10)	65 (8)	56 (6)	16 (4)	16 (3)	22 (3)
Older male	37 (9)	42 (7)	43 (7)	14 (5)	14 (5)	21 (5)

Table 2. VOT means and standard deviations (ms) calculated over talker-specific means for each stop category and social group.

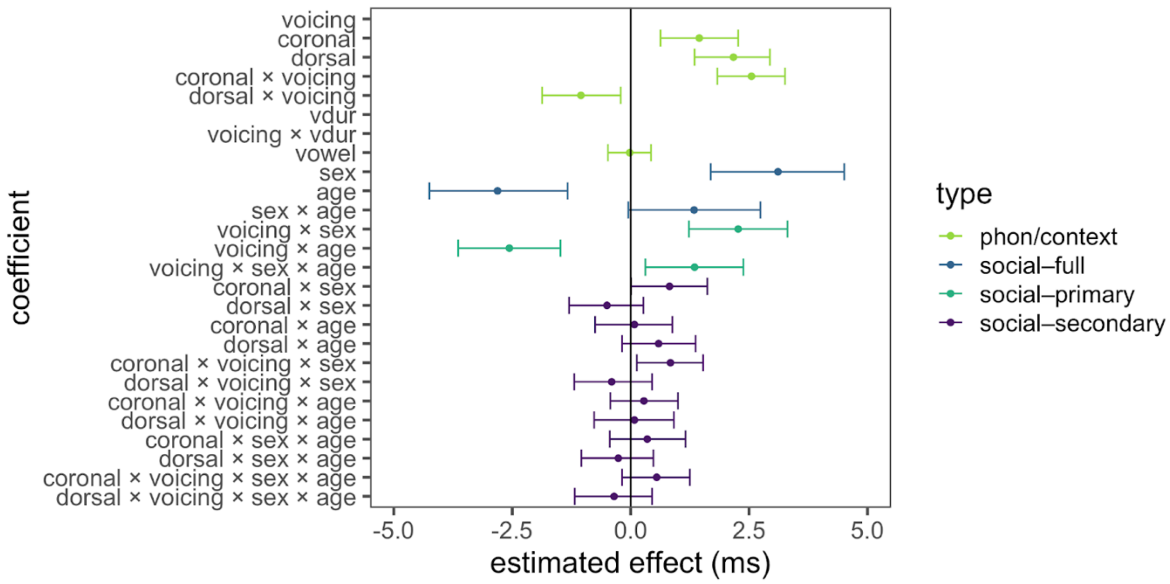

Figure 3. Mean estimated beta coefficient on VOT with bars extending over the 95% credible interval. Each beta coefficient can be interpreted as the deviation in milliseconds from the mean VOT in the dataset (centred at 0 ms). Effects are colour-coded as a phonological or contextual influence, a social influence on the VOT of the full class of stops, a social influence on the primary phonological correlate of VOT, or a social influence interacting with a secondary phonological correlate of VOT. Note that the effects of voice, vowel duration, and their interaction were too large to display at this scale.

4.2.2. COG

Variation in COG by sex, age, and stop category is shown in Figure 4 and Table 3. The model revealed credible influences of several linguistic and social factors on COG variation (see Figure 5). As expected, coronal stops were approximately 614 Hz higher than the average COG, and dorsal stops approximately 323 Hz lower than the average COG (coronal: 614.41, 95% CI [538.06, 692.65]; dorsal: -322.72, 95% CI [-395.06, -251.31]). In addition, voiceless stops had higher COGs than voiced COGs (voice: 609.77, 95% CI [546.97, 673.49]); however, this was reliably modulated by place of articulation: the voicing contrast in COG was reliably enhanced between coronal stops and diminished between dorsal stops (coronal × voice: 350.11, 95% CI [287.95, 409.82]; dorsal × voice: -206.68, 95% CI [-267.53, -145.90]). The main effect of voicing could potentially reflect a minor influence of vocal fold vibration on voiced stops and/or increased airflow associated with voiceless stops (e.g. Zue 1976; Chodroff and Wilson 2014). Stop COG was also reliably higher before high vowels than non-high vowels (vowel: 59.35, 95% CI [32.66, 86.46]), as well as before longer vowels (i.e. slower speaking rates; vdur: 410.57, 95% CI [243.87, 574.50])

Among social factors, sex and age reliably structured COG variation, largely paralleling the VOT patterns: females had higher COGs than males, and younger speakers had higher COGs than older speakers (sex: 162.69, 95% CI [61.22, 260.18]; age: -235.43, 95% CI [-335.41, -134.82]). In addition, older men had overall lower COGs relative to what would be expected based on the main effects of sex and age alone (sex × age: 96.98, 95% CI [1.60, 190.37]). Social factors also had large interactions with place of articulation: coronal stop COGs were particularly high across female speakers (coronal × sex: 151.76, 95% CI [72.08, 230.62]), and markedly low across older speakers (coronal × age: -104.32, 95% CI [-178.76, -28.73]). Dorsal COGs were also somewhat higher across older speakers than would be expected based on the other effects (dorsal × age: 110.28, 95% CI [41.88, 176.89]). Finally, older speakers had a greater contrast in COG between voiced and voiceless stops relative to younger speakers (voicing × age: 63.61, 95% CI [0.97, 125.61]), except between [tʰ] and [d] (coronal × voicing × age: -67.93, 95% CI [-128.06, -5.66]). No other interactions were reliable in their direction.

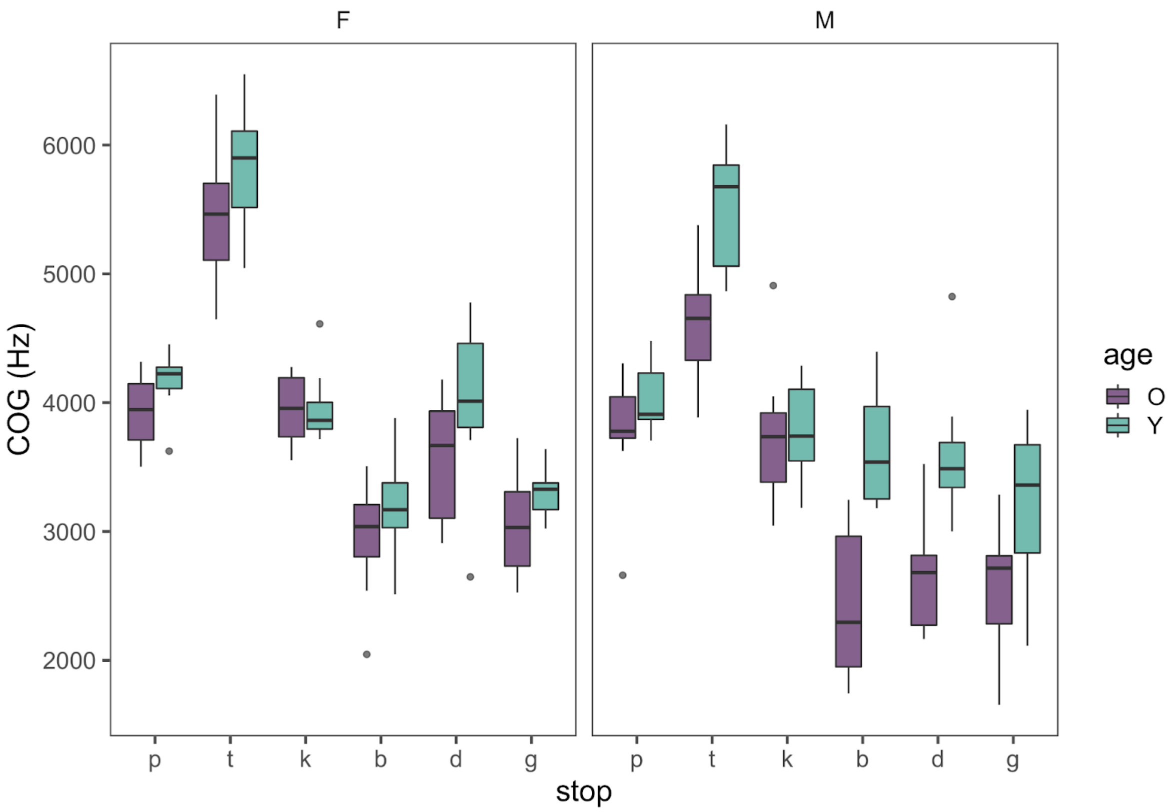

Figure 4. Variation in COG (Hz) for [pʰ tʰ kʰ b d g] across sex and age.

COG	[pʰ]	[tʰ]	[kʰ]	[b]	[d]	[g]
Younger female	4167 (254)	5836 (467)	3969 (298)	3211 (418)	4013 (678)	3295 (200)
Younger male	4030 (280)	5521 (489)	3779 (410)	3637 (456)	3622 (549)	3203 (623)
Older female	3929 (296)	5454 (493)	3954 (262)	2958 (422)	3551 (468)	3078 (427)
Older male	3793 (485)	4589 (451)	3733 (552)	2417 (551)	2689 (487)	2564 (504)

Table 3. COG means and standard deviations (Hz) calculated over talker-specific means for each stop category and social group.

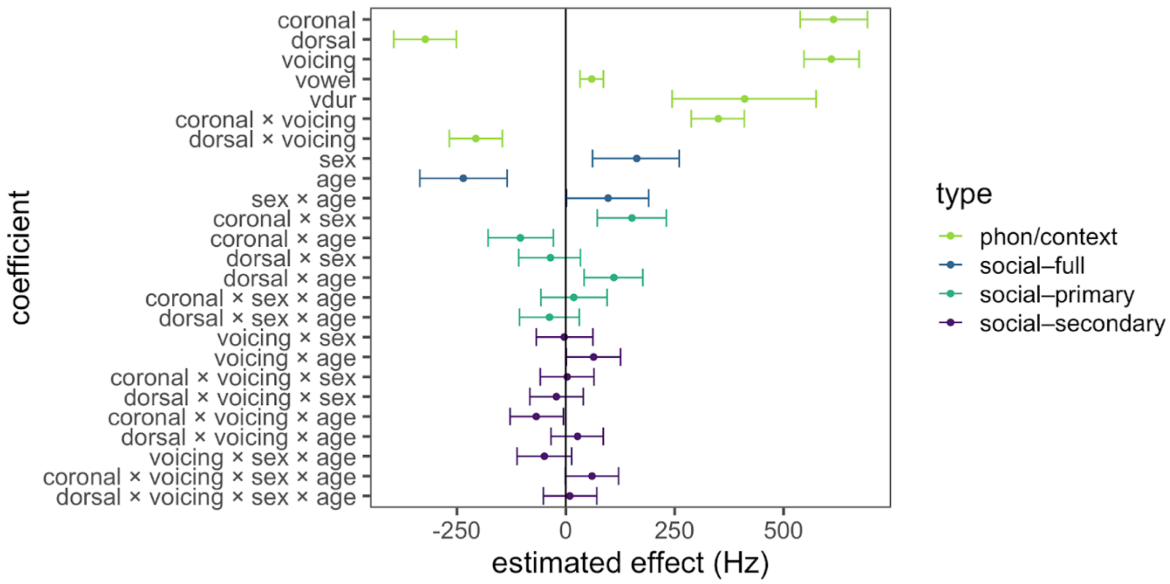

Figure 5. Mean estimated beta coefficient on COG with bars extending over the 95% credible interval. Each beta coefficient can be interpreted as the deviation in Hertz from the mean COG in the dataset (centred at 0 Hz). Effects are colour-coded as a phonological or contextual influence, a social influence on the COG of the full class of stops, a social influence on the primary phonological correlate to COG, or a social influence interacting with a secondary phonological correlate to COG.

4.3. Discussion

Derby stop realisation contains rich phonological, contextual, and - critically - social variation. Many of these effects were fully in line with expectations. Main influences on VOT were phonological voicing, following vowel duration (as a proxy for speaking rate), and the interaction between voicing and vowel duration. The interaction between place of articulation and voicing had additional reliable influences, with VOT increasing with more posterior places of articulation. In contrast to previous studies of English, following vowel height did not reliably influence VOT (cf. Nearey and Rochet, 1994). For COG, place of articulation had the largest influence on variation. Phonological voicing had a similarly large effect on realisation within each place of articulation, and the interactions between place and voicing were also reliable. Notably, dorsal voiced and voiceless stops did not differ to the same degree as coronal and labial stop contrasts. This minimised effect of voicing on dorsal stops has been observed previously (e.g. Chodroff and Wilson 2014). COG was also reliably raised before high vowels, as well as before longer vowels.

With respect to our hypothesis regarding social influences, social factors largely modulated the phonetic realisation of the full class of stops, or the primary phonetic targets of a phonological feature. Specifically, of all social influences, the largest beta estimates were on the overall realisation of stop VOT and COG. This suggests that the primary target of social influence is on the full natural class of stop consonants. The second largest modulations involving social factors targeted the magnitude of the phonological contrast primarily conveyed by the relevant phonetic dimension. The interactions between social factors and non-primary contrasts were minimal.

For VOT, female speakers had longer VOTs than male speakers, and older speakers had shorter VOTs than younger speakers. These patterns are in keeping with previous studies on

English (see section 2.4; e.g. Swartz 1992; Ryalls et al. 1997; Whiteside et al. 2004; Ma et al. 2018). Social factors also interacted reliably with the voicing contrast (i.e. the sex x voicing and age x voicing interactions). These indicated females had an overall greater voicing contrast using VOT (and higher voiceless stop VOTs) than males; in addition, older speakers had a smaller voicing contrast than younger speakers. Minor but reliable influences of social factors were also found on the size of the voicing contrast specifically for coronal segments. Critically, interactions between social factors and place of articulation alone were again minimal, and mostly not reliable in their direction. The one exception was the finding that females had slightly longer VOTs for coronal stops relative to other places of articulation compared to males. Indeed, VOT primarily reflected the voicing contrast, and the largest variation in social influence was in this particular contrast. Social factors had negligible influence on contrasts between places of articulation.

Female speakers had higher COGs than male speakers, younger speakers had higher COGs than older speakers, and older men also had particularly low COGs. The low COGs observed among older speakers could reflect multiple sources: COG does primarily capture information regarding place, and therefore the lowered COG could reflect more retracted articulations for dorsals and especially coronals (the latter in line with impressionistic observation; Docherty and Foulkes 1999: 51). However, it is also not immune to influences from phonetic voicing: it is possible that older speakers may also employ a higher degree of phonetic voicing among voiced stops than younger speakers. It is also possible that the lowered COG reflects a decreased lung volume among older speakers (Mead et al. 1967). Isolating these exact influences on COG remains for further research.

COG serves as a primary acoustic-phonetic correlate of place of articulation contrasts and only a secondary correlate of voicing; as such, the prediction is for interactions between social factors and place of articulation to have a greater influence on COG than interactions between social factors and voicing. Results largely confirmed these predictions. The largest social effects on COG were on the entire natural class of stops. The second largest set of social influences were with the place of articulation contrast (i.e. coronal x sex, coronal x age, dorsal x age). Interestingly, despite the fact that COG is only a secondary correlate of the voicing contrast, a reliable modulation was also observed in the interaction between age and voicing, as well as the three-way interaction between age, coronal place, and voicing. Specifically, older speakers had a greater voicing contrast in COG relative to younger speakers, but this effect was somewhat reduced for the voicing contrast in coronals. Nevertheless, the largest reliable interactions were either between social factors and stops as a whole class, or between social factors and effects of place of articulation. Given these patterns, we infer that social factors have a particular respect for linguistic structure and the phonetic realisation of that structure.

5. Conclusion

In this chapter, we have reviewed variation in stop realisation relating to linguistic, biological, and social factors. Social variation in stops is present along each of the phases of stop articulation including the closing phase, hold phase, and release phase. In addition, considerable work has also investigated and identified variation in the coupling between the oral and laryngeal gestures, which has most commonly been studied in terms of VOT. Regional and/or social variation is observed in several acoustic and articulatory correlates of stop production. Stops are thus potentially a very rich resource for social variation. However, the review in section 3 reveals clear gaps that future research could easily fill.

VOT is one of the most widely studied phonetic dimensions of all segments. There is certainly evidence for social and individual variation in VOT, but the actual effects tend to be small and inconsistent. It is unclear how useful these are in perception: for example, even where small sex effects are observed, as far as we are aware no study has shown that listeners orient to these differences as a marker for gender. We can only speculate as to why this might be, but VOT has high variability both within and across speakers. As suggested by Kleinschmidt (2019), this contributes to it having less utility for determining social categories relative to other phonetic measures such as vowel formants. In addition, he shows that social categories are less informative for determining the identity of a stop given its VOT than the identity of a vowel given its formants. We might further ask whether this point is a more general one, namely that social marking is more successful when applied to frequency-based features or discrete segmental units rather than durational features (we noted in section 3.2, for example, that there seems little evidence for social variation in the hold phase of stops—another temporal feature—either). Variation in the temporal domain certainly does contribute socioindexical information, for example, via speech rate, rhythm, and stress. VOT might therefore contribute to variation at this level as one of many features that can vary temporally, but presumably it is less free to carry its own social value.

Moreover, VOT is not the only dimension of stop variation. As demonstrated above, stops also vary along several other dimensions. The increasing availability of large-scale speech corpora and automatic analysis tools offers tremendous scope to explore sociophonetic variation in the acoustic form of stops. Our case study in section 4, which drew on a fairly small corpus, illustrates one such approach. Further research can also draw on perceptual and articulatory measures for understanding variation in stops. Finally, as with many other sociophonetic studies, analysis of stops enables us to explore theoretical questions of fundamental importance: the structure of the lexicon, speech production and perception, and the cognitive representation of socio-indexical information.

6. References

Abramson, A. S., & Whalen, D. H. (2017) Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, *63*, 75-86.

Abrego-Collier, C., Grove, J., Sonderegger, M., & Alan, C. L. (2011) Effects of speaker evaluation on phonetic convergence. In W.-S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 192-195). City University of Hong Kong.

Alexander, K. (2014) *Perception of the boundary between singleton and geminate plosives by Greek Cypriots: a sociophonetic perspective.* PhD thesis, University of York.

Allen, J. S., & Miller, J. L. (1999) Effects of syllable-initial voicing and speaking rate on the temporal characteristics of monosyllabic words. *Journal of the Acoustical Society of America* 106 (4) : 2031. DOI: http://dx.doi.org/10.1121/1.427949

Allen, J. S., Miller, J. L., & DeSteno, D. (2003) Individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America*, *113*(1), 544–552. https://doi.org/10.1121/1.1528172

Al-Tamimi, J., & Khattab, G. (2015) Acoustic cue weighting in the singleton vs geminate contrast in Lebanese Arabic: The case of fricative consonants. The Journal of the Acoustical Society of America, 138(1), 344-360.

Al-Tamimi, J., & Khattab, G. (2018) Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics*, *71*, 306-325.

Andruski, J. E., Blumstein, S. E., & Burton, M. (1994) The effect of subphonetic differences on lexical access. *Cognition*, 52(3), 163-187.

Aoyama, K., & Reid, L. (2006) Cross-linguistic tendencies and durational contrasts in geminate consonants: An examination of Guinaang Bontok geminates. Journal of the International Phonetic Association, 36(2), 145-157. doi:10.1017/S0025100306002520

Arvaniti, A. (2007) Greek Phonetics: The State of the Art. Journal of Greek Linguistics 8, 97–208. doi 10.1075/jgl.8.08arv

Avery, P., & Rice, K. (1989) Segment structure and coronal underspecification. *Phonology*, *6*(2), 179-200.

Backley, P. (1993). Coronal: the undesirable element. *UCL Working Papers in Linguistics*, 5, 301-323.

Baese-Berk, M., & Goldrick, M. (2009) Mechanisms of interaction in speech production. *Language and cognitive processes*, *24*(4), 527-554.

Baran, J. A., Laufer, M. Z., & Daniloff, R. (1977) Phonological contrastivity in conversation: A comparative study of voice onset time. *Journal of Phonetics*, *5*(4), 339-350.

Baranowski, M., & Turton, D. (2020) TDd-deletion in British English: New evidence for the long-lost morphological effect. *Language Variation and Change, 32*(1), 1-23. doi:10.1017/S0954394520000034

Bates, D., Sarkar, D., Bates, M. D., & Matrix, L. (2007) The lme4 package. *R package version*, *2*(1), 74.

Baum, S. R., & Ryan, L. (1993) Rate of speech effects in aphasia: Voice onset time. *Brain and Language*, *44*(4), 431-445.

Bayley, R. (1994) Consonant cluster reduction in Tejano English. *Language Variation and Change, 6*(3), 303-326. doi:10.1017/S0954394500001708

Becker, M., Ketrez, N., & Nevins, A. (2011) The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. Language, 84-125.

Beckman, J., Jessen, M., & Ringen, C. (2013). Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics*, *49*(2), 259-284.

Benjamin, B. J. (1982) Phonological performance in gerontological speech. *Journal of Psycholinguistic Research*, *11*(2), 159–167. https://doi.org/10.1007/BF01068218

Benjamin, B. (1997) Speech production of normally aging adults. *Seminars in speech and language 18*, 135-141.

Benor, Sarah. (2001) The learned /t/: Phonological variation in Orthodox Jewish English. Penn Working Papers in Linguistics (Selected Papers from New Ways of Analyzing Variation 29) 7:1–16.

Berry, J., & Moyle, M. (2011) Covariation among vowel height effects on acoustic measures. *The Journal of the Acoustical Society of America*, *130*(5), EL365-EL371.

Blumstein, S. E., Myers, E. B., & Rissman, J. (2005) The perception of voice onset time: an fMRI investigation of phonetic category structure. *Journal of Cognitive Neuroscience*, *17*(9), 1353-1366.

Blumstein, S. E., & Stevens, K. N. (1979) Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *The Journal of the Acoustical Society of America*, *66*(4), 1001-1017.

Blumstein, S. E., & Stevens, K. N. (1980) Perceptual invariance and onset spectra for stop consonants in different vowel environments. *The Journal of the Acoustical Society of America*, *67*(2), 648-662.

Boario, A. (2009) Il raddoppiamento fonosintattico nelle varietà di parlanti adolescenti nativi e non nativi. In C. Consani, P. Desideri, F. Guazzelli & C. Perta (Eds.), Alloglossie e comunità alloglotte nell'Italia contemporanea. Teorie, applicazioni e descrizioni, prospettive (pp. 383-398) Roma: Bulzoni.

Bonneau, A., Djezzar, L., & Laprie, Y. (1996) Perception of the place of articulation of French stop bursts. *The Journal of the Acoustical Society of America*, *100*(1), 555–564.

Bowerman, S. (2008) White South African English: phonology. In R. Mesthrie (ed.) Varieties of English volume 4. Africa, South and Southeast Asia. Berlin & New York: Mouton de Gruyter. pp. 164-176.

Brandt, E., & Simpson, A. P. (2021) The production of ejectives in German and Georgian. *Journal of Phonetics*, *89*, 101111.

Bucholtz, M. (1996) Geek the girl: Language, femininity, and female nerds. In Warner, N., Ahlers, J., Bilmes, L., Oliver, M., Wertheim, S., and Chen, M. (Eds.) *Gender and Belief Systems: Proceedings of the Fourth Berkeley Women and Language Conference.* Berkeley: Berkeley Women and Language Group. pp. 119-131.

Bundgaard-Nielsen, R. L., & O'Shannessy, C. (2021) When more is more: The mixed language Light Warlpiri amalgamates source language phonologies to form a near-maximal inventory. Journal of Phonetics, 85, 101037.

Bürkner, P.-C. (2018) Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, *10*(1), 395–411. https://doi.org/10.32614/rj-2018-017

Byrd, D. (1993) 54,000 American stops. *UCLA Working Papers in Phonetics*, *83*, 97–116.

Caramazza, A., & Yeni-Komshian, G. H. (1974) Voice onset time in two French dialects. Journal of phonetics, 2(3), 239-245.

Cassidy, S., & Harrington, J. (1995) The Place of Articulation Distinction in Voiced Oral Stops: Evidence from Burst Spectra and Formant Transitions. *Phonetica*, *52*, 263–284.

Cho, T., Jun, S., & Ladefoged, P. (2002) Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, *30*, 193–228.

Cho, T., & Keating, P. A. (2001) Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, *29*(2), 155–190. https://doi.org/10.1006/jpho.2001.0131

Cho, T., & Keating, P. A. (2009) Effects of initial position versus prominence in English. *Journal of Phonetics*, *37*(4), 466–485. https://doi.org/10.1016/j.wocn.2009.08.001

Cho, T., & Ladefoged, P. (1999) Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27 (2) : 207. DOI: http://dx.doi.org/10.1006/jpho.1999.0094

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

Cho, T., Whalen, D. H., & Docherty, G. (2019) Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. *Journal of Phonetics, 72*, 52-65.

Chodroff, E., Golden, A., & Wilson, C. (2019) Covariation of stop voice onset time across languages: Evidence for a universal constraint on phonetic realization. *The Journal of the Acoustical Society of America*, *145*(1), EL109–EL115. https://doi.org/10.1121/1.5088035

Chodroff, E., & Wilson, C. (2014) Burst spectrum as a cue for the stop voicing contrast in American English. *The Journal of the Acoustical Society of America*, *136*(5), 2762–2772. https://doi.org/Doi 10.1121/1.4896470

Chodroff, E., & Wilson, C. (2017) Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. *Journal of Phonetics*, *61*, 30–47. https://doi.org/10.1016/j.wocn.2017.01.001

Chodroff, E., & Wilson, C. (2018) Predictability of stop consonant phonetics across talkers: Between-category and within-category dependencies among cues for place and voice. *Linguistics Vanguard*, *4*(s2) https://doi.org/10.1515/lingvan-2017-0047

Choi, H. (2003) Prosody-induced acoustic variation in English stop consonants. In M. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2661–2664).

Clark, L., & Watson, K. (2016) Phonological leveling, diffusion, and divergence:/t/lenition in Liverpool and its hinterland. *Language Variation and Change*, *28*(1), 31-62.

Clayton, I. (2017) Preaspiration in Hebrides English. Journal of the International Phonetic Association, 47(2), 155-181. doi:10.1017/S002510031700007X

Clothier, J. & Loakes, D. (2018) Coronal Stop VOT in Australian English: Lebanese Australians and Mainstream Australian English. Epps, J., Wolfe, J., Smith, J., Jones, C. (Eds.), pp.13-16. Australasian Speech Science and Technology Australia (ASSTA).

Clumeck, H., Barton, D., Macken, M. A., & Huntington, D. A. (1981) The aspiration contrast in Cantonese word-initial stops: Data from children and adults. *Journal of Chinese Linguistics*, *9*(2), 210–225.

Cohen Priva, U. (2017) Informativity and the actuation of lenition. Language 93(3), 569–597. doi:10.1353/lan.2017.0037.

Cohn, A. C. (1993) Nasalisation in English: Phonology or phonetits. *Phonology*, *10*(1), 43–81. https://doi.org/10.1017/S0952675700001731

Cole, J. S., Choi, H., Kim, H., & Hasegawa-Johnson, M. (2003) The effect of accent on the acoustic cues to stop voicing in Radio News speech. In M. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 15–18).

Cole, J., Kim, H., Choi, H., & Hasegawa-Johnson, M. (2007) Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech. *Journal of Phonetics*, 35(2), 180-209.

Cotter, W. (2014) (q) as a sociolinguistic variable in the Arabic of Gaza City. In Perspectives on Arabic Linguistics XXVIII. Papers from the Annual Symposium on Arabic Linguistics, Gainesville, Florida, 2014. (pp. 229-246) John Benjamins.

Davidson, L. (2011) Characteristics of stop releases in American English spontaneous speech. *Speech Communication*, *53*(8), 1042-1058.

Davidson, L. (2016) Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics*, *54*, 35-50.

Davidson, L. (2021) Effects of word position and flanking vowel on the implementation of glottal stop: Evidence from Hawaiian. *Journal of Phonetics*, 88, 101075.

DiCanio, C. T. (2012) The phonetics of fortis and lenis consonants in Itunyoso Trique. *International Journal of American Linguistics*, *78*(2), 239–272.

Docherty, G.J. (1992) *The timing of voicing in British English obstruents*. Walter de Gruyter.

Docherty, G.J. & Foulkes, P. (1999) Newcastle upon Tyne and Derby: instrumental phonetics and variationist studies. In Foulkes, P. & Docherty, G.J. (Eds.) *Urban Voices: Accent Studies in the British Isles*. London: Arnold. pp. 47-71.

Docherty, G.J. & Foulkes, P. (2005) Glottal variants of (t) in the Tyneside variety of English: an acoustic profiling study. In Hardcastle, W. & Mackenzie Beck, J. (Eds.) *A Figure of Speech – a Festschrift for John Laver*. London: Lawrence Erlbaum. pp. 173-199.

Docherty, G.J. & Foulkes, P. (2014) An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua* 142, 42-56.

Docherty, G.J., Foulkes, P., Milroy, J., Milroy, L. & Walshaw, D. (1997) Descriptive adequacy in phonology: a variationist perspective. *Journal of Linguistics* 33, 275-310.

Docherty, G.J., Watt, D., Llamas, C., Hall, D., & Nycz, J. (2011) Variation in voice onset time along the Scottish-English Border. In W.-S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 591–594). City University of Hong Kong.

Duran, D., Bruni, J., Dogil, G., & Roux, J. (2017) The social life of Setswana ejectives. In *INTERSPEECH* (pp. 3787-3791).

Ellingson Eddington, D., & Brown, E.K. (2021) A production and perception study of /t/-glottalization and oral releases following glottals in the United States. *American Speech 96* (1): 78–104. doi: https://doi.org/10.1215/00031283-8620501

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971) Speech perception in infants. *Science*, *171*(3968), 303-306.

Elías-Ulloa, J. (2004). Quantity (in) sensitivity and underlying glottal-stop deletion in Capanahua. *Coyote Papers* 13. http://hdl.handle.net/10150/126614

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

Fabricius, A. (2002) Ongoing change in modern RP: Evidence for the disappearing stigma of t-glottalling. *English world-wide*, *23*(1), 115-136.

Falahati, R. (2013) *Gradient and Categorical Consonant Cluster Simplification in Persian: An Ultrasound and Acoustic Study*. PhD thesis, University of Ottawa.

Fant, G. (1973) *Speech Sounds and Features.* The MIT Press.

Fimone, W. (2020) Onset glottal stop deletion in Suva Rotuman. *Asia-Pacific Language Variation*, *6*(2), 196-221.

Finn, P. (2008) Cape Flats English: phonology. In R. Mesthrie (ed.) Varieties of English volume 4. Africa, South and Southeast Asia. Berlin & New York: Mouton de Gruyter. pp. 200-222.

Fischer-Jørgensen, E. (1954) Acoustic analysis of stop consonants. *Le Maître Phonétique*, *32*, 42-59.

Fish, M. S., García-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. K. (2017) Infant-directed speech in English and Spanish: Assessments of monolingual and bilingual caregiver VOT. *Journal of Phonetics*, *63*, 19-34.

Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988) Statistical analysis of word-initial voiceless obstruents: Preliminary data. *The Journal of the Acoustical Society of America*, *84*(1), 115–123. https://doi.org/10.1121/1.396977

Foulkes, P. (2010) Exploring social-indexical variation: a long past but a short history. *Laboratory Phonology* 1, 5-39.

Foulkes, P. & Docherty, G.J. (2006) The social life of phonetics and phonology. *Journal of Phonetics* 34, 409-438.

Foulkes, P., Docherty, G.J., & Jones, M.J. (2010) Analysing stops. In Di Paolo, M. & Yaeger-Dror, M. *Sociophonetics: A Student's Guide*. London: Routledge. pp. 58-71.

Foulkes, P., Docherty, G.J., & Watt, D. (2005) Phonological variation in child directed speech. *Language* 81, 177-206.

Foulkes, P., & Hay, J. (2015) The emergence of sociophonetic structure. In B. MacWhinney & W. O'Grady (Eds.) *The Handbook of Language Emergence*. Oxford: Blackwell. pp. 292-313.

Foulkes, P., & Hughes, V. (in press) Dialectological and sociolinguistic foundations of forensic speaker comparison. To appear in F. Nolan, K. McDougall & T. Hudson (Eds.), *Oxford Handbook of Forensic Phonetics*. Oxford: Oxford University Press.

Fourakis, M., & Port, R. (1986) Stop epenthesis in English. *Journal of Phonetics* 14, 197–221.

Fox, N. P., Reilly, M., & Blumstein, S. E. (2015) Phonological neighborhood competition affects spoken word production irrespective of sentential context. *Journal of Memory and Language*, *83*, 97-117.

Fox, S. & Torgersen, E. (2018) Language change and innovation in London: Multicultural London English. In N. Braber & S. Jansen (Eds.), *Sociolinguistics in England* (pp. 189-213). Palgrave Macmillan, London.

Fromont, R. A., & Hay, J. (2012) LaBB-CAT: An annotation store. Proceedings of the 14th Australasian International Conference on Speech Science and Technology, 113−117.

Fruehwald, J. (2017) The role of phonology in phonetic change. *Annual Review of Linguistics*, *3*, 25–42. https://doi.org/10.1146/annurev-linguistics-011516-034101

Gadet, F., & Hambye, P. (2014) Contact and Ethnicity in 'Youth Language' Description: In Search of Specificity. In R. Nicolaï (ed.) Questioning Language Contact: Limits of Contact, Contact at Its Limits, 183–217. Leiden: Brill.

Gandour, J., Petty, S. H., Dardarananda, R., Dechongkit, S., & Mukngoen, S. (1986) The acquisition of the voicing contrast in Thai: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*, *13*(3), 561–572. https://doi.org/10.1017/S0305000900006887

Gao, J., & Arai, T. (2019) Plosive (de-) voicing and f0 perturbations in Tokyo Japanese: Positional variation, cue enhancement, and contrast recovery. *Journal of Phonetics*, *77*, 100932.

Ganong, W. F. (1980) Phonetic categorization in auditory word perception. *Journal of experimental psychology: Human Perception and Performance*, *6*(1), 110.

Gargesh, R. (2008) Indian English: phonology. In R. Mesthrie (ed.) Varieties of English volume 4. Africa, South and Southeast Asia. Berlin & New York: Mouton de Gruyter. pp. 231-24376.

Gorham-Rowan, M.M., & Laures-Gore, J. (2006) Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of communication disorders*, *39*(3), 171-184.

Gósy, M., & Ringen, C. O. (2009) Everything you always wanted to know about VOT in Hungarian. In *IXth International Conference on the Structure of Hungarian*.

Gráczi, T. E., & Kohári, A. (2014) Multiple bursts in Hungarian voiceless plosives and VOT measurements. In Proceedings of the 10th International Seminar on Speech Production (ISSP), Cologne, Germany (pp. 158-161).

Gurevich, N. 2004. *Lenition and Contrast: the Functional Consequences of Certain Phonetically Conditioned Sound Changes*. Routledge: New York and London.

Guy, G. R., & Hinskens, F. (2016) Linguistic coherence: Systems, repertoires and speech communities. *Lingua*, *172–173*, 1–9. https://doi.org/10.1016/j.lingua.2016.01.001

Hale, M., & Reiss, C. (2000) " Substance abuse" and" dysfunctionalism": Current trends in phonology. *Linguistic Inquiry*, *31*(1), 157-169.

Hall, T. A., & Hamann, S. (2006) Towards a typology of stop assibilation. *Linguistics*, 44, 1195-1236.

Halle, M., Hughes, G. W., & Radley, J. A. (1957) Acoustic properties of stop consonants. *Journal of Acoustical Society of America*, *29*(1), 107–116.

Ham, W. 2001. *Phonetic and Phonological Aspects of Geminate Timing*. New York: Routledge.

Hanson, H. M., & Stevens, K. N. (2003) Models of aspirated stops in English. In M. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 783–786).

Harrington, J. (2010) *The Phonetic Analysis of Speech Corpora*. 292. Malden, MA: Wiley-Blackwell.

Hay, J. & Foulkes, P. (2016) The evolution of medial (-t-) over real and remembered time. *Language* 92(2): 298-330.

Hejná, M. 2015. *Pre-aspiration in Welsh English: A Case Study of Aberystwyth*. PhD thesis, University of Manchester.

Helgason, P. (2002) *Preaspiration in the Nordic Languages. Synchronic and Diachronic Aspects*. PhD thesis, Stockholm University. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A193411&dswid=-4337

Helgason, P. & Ringen, C. (2008) Voicing and aspiration in Swedish stops. *Journal of Phonetics, 36*(4), 607-628.

Herd, W. (2020) Sociophonetic voice onset time variation in Mississippi English. *The Journal of the Acoustical Society of America*, *147*(1), 596-605.

Hickey, R. (2017) Twentieth-Century Received Pronunciation: Stop Articulation. In Hickey, R. (ed.) *Listening to the Past.* Cambridge: CUP. pp. 66-84.

Hombert, J. M., Ohala, J. J., & Ewan, W. G. (1979) Phonetic explanations for the development of tones. *Language*, 37-58.

Hualde, J. I. (2005) *The Sounds of Spanish*. Cambridge University Press

Jamin, M. (2004) 'Beurs' and accent des cités: A case study of linguistic diffusion in La Courneuve. *Contemporary French and Francophone Studies* 8 (2): 169–76.

Kang, K. H. (2010) Generational differences in the perception of Korean stops. *Phonetics and Speech Sciences*, *2*(3), 3-10.

Kang, K. H., & Guion, S. G. (2008) Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America*, *124*(6), 3909-3917.

Keating, P. A. (1984) Phonetic and phonological representation of stop consonant voicing. *Language*, *60*(2), 286–319

Keating, P. A. (1990) Phonetic representations in a generative grammar. *Journal of Phonetics*, *18*(3), 321–334.

Keating, P., Linker, W., & Huffman, M. (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics*, *11*(3), 277-290.

Keating, P. A., Mikoś, M. J., & Ganong III, W. F. (1981) A cross-language study of range of voice onset time in the perception of initial stop voicing. *The Journal of the Acoustical Society of America*, *70*(5), 1261-1271.

Keshet, J., Sonderegger, M., Knowles, T. (2014) AutoVOT: A tool for automatic measurement of voice onset time using discriminative structured prediction [Computer program]. Version 0.91, retrieved July 2019 from https://github.com/mlml/autovot/.

Kessinger, R. H., & Blumstein, S. E. (1997) Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics*, *25*(2), 143–168. https://doi.org/10.1006/jpho.1996.0039

Kessinger, R. H., & Blumstein, S. E. (1998) Effects of speaking rate on voice-onset time and vowel production: Some implications for perception studies. *Journal of Phonetics*, *26*(2), 117–128. https://doi.org/10.1006/jpho.1997.0069

Kewley-Port, D. (1982) Measurement of formant transitions in naturally produced stop consonant–vowel syllables. *The Journal of the Acoustical Society of America*, *72*(2), 379-389.

Kewley-Port, D. (1983) Time-varying features as correlates of place of articulation in stop consonants. *Journal of Acoustical Society of America*, *73*(1), 322–335.

Khattab, G. (2003) Age, input, and language mode factors in the acquisition of VOT by English-Arabic bilingual children. In M. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences Proceedings* (pp. 3213-3216).

Khattab, G. and Al-Tamimi, J. (2014) Geminate timing in Lebanese Arabic: the relationship between phonetic timing and phonological structure. *Laboratory Phonology, 5*(2), 231-269. https://doi.org/10.1515/lp-2014-0009

Khattab, G. & Foulkes, P. (this volume) Arabic sociophonetics. In Strelluf, C. (ed.) *The Routledge Handbook of Sociophonetics*. Routledge.

Khattab, G., Al-Tamimi, F., & Heselwood, B. (2006) Acoustic and auditory differences in the /t/-/T/ opposition in male and female speakers of Jordanian Arabic. In S. Boudelaa (ed.) Perspectives on Arabic Linguistics XVI. Cambridge, UK: John Benjamins. pp. 131-160.

Kirchner, R. (2001) *An Effort Based Approach to Consonant Lenition*. Routledge.

Kirkham, S. (2011) The acoustics of coronal stops in British Asian English. In W.-S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 1102–1105). City University of Hong Kong.

Kirov, C., & Wilson, C. (2012) The specificity of online variation in speech production. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 587–592).

Klatt, D. H. (1975) Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research* 18 (4) : 686. DOI: http://dx.doi.org/10.1044/jshr.1804.686

Kleber, F. (2018) VOT or quantity: What matters more for the voicing contrast in German regional varieties? Results from apparent-time analyses. Journal of Phonetics, 71, 468-486.

Kleinschmidt, D. F., Weatherholtz, K., & Jaeger, T. F. (2018) Sociolinguistic perception as inference under uncertainty. *Topics in Cognitive Science*, *10*(4), 818-834.

Kochetov, A., & Pouplier, M. (2008) Phonetic variability and grammatical knowledge: An articulatory study of Korean place assimilation. *Phonology*, *25*(3), 399-431.

Koenig, L. (2001) Distributional characteristics of VOT in children's voiceless aspirated stops and interpretation of developmental trends. J. Speech Lang. Hear. Res., 44, 1058-1068.

Kohler, K. J. (1994) Glottal stops and glottalization in German. *Phonetica*, *51*(1-3), 38-51.

Kong, E. J., Holliday, J. J., & Lee, H. (2022). Post-adolescent changes in the perception of regional sub-phonemic variation. *Journal of Phonetics*, *90*, 101114.

Kozminska, K. (2015) Preliminary results of a sociophonetic study of VOT and Polish transnational identities in the UK. In The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: the University of Glasgow. ISBN 978-0-85261-941-4. Paper number 0827.

Kuang, J. (2019) Contextual variation of glottalic stops in Q'anjob'al. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia 2019 (pp. 1114-1118). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Kubozono, H. (Ed.) (2017) *The Phonetics and Phonology of Geminate Consonants*. Oxford University Press.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017) lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1-26.

Labov, W. (1972) *Sociolinguistic patterns*. Oxford: Blackwell.

Labov, W. (2006) A sociolinguistic perspective on sociophonetic research. *Journal of Phonetics*, *34*(4), 500-515.

Ladefoged, P. (2005) *Vowels and Consonants* (2nd edition). Malden MA: Blackwell.

Ladefoged, P. & Maddieson, I. (1996) The Sounds of the World's Languages. Oxford: Blackwell.

Lass, R. (1984) *Phonology: An introduction to basic concepts*. Cambridge University Press.

Lavoie, L. M. (2001) *Consonant Strength: Phonological Patterns and Phonetic Manifestations*. Garland Publishing, Inc.

Lee, H., Holliday, J. J., & Kong, E. J. (2020) Diachronic change and synchronic variation in the Korean stop laryngeal contrast. *Language and Linguistics Compass*, 14(7), e12374.

Lee, H., & Jongman, A. (2012) Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association*, *42*(2), 145-169.

Lee, H., Politzer-Ahles, S., & Jongman, A. (2013) Speakers of tonal and non-tonal Korean dialects use different cue weightings in the perception of the three-way laryngeal stop contrast. *Journal of Phonetics*, *41*(2), 117-132.

Levon, E. (2006) Hearing "gay": Prosody, interpretation, and the affective judgments of men's speech. *American speech*, *81*(1), 56-78.

Li, F. (2013) The effect of speakers' sex on voice onset time in Mandarin stops. *The Journal of the Acoustical Society of America*, *133*(2), EL142-7. https://doi.org/10.1121/1.4778281

Lisker, L. (1986) "Voicing" in English: A Catalogue of Acoustic Features Signaling /b/ Versus /p/ in Trochees. *Language and Speech*, *29*(1), 3–11.

Lisker, L., & Abramson, A. S. (1964) A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20(3), 384-422.

Lisker, L., & Abramson, A. S. (1967) Some effects of context on voice onset time in English stops. *Language and Speech* 10(1), 1–28.

Loakes, D., McDougall, K., Clothier, J., Hajek, J., & Fletcher, J. (2018) Sociophonetic variability of post-vocalic/t/in Aboriginal and mainstream Australian English. In *Proceedings of the 17th Speech Science and Technology Conference (SST2018)* (pp. 1-4) ASSTA: Coogee, Sydney.

Löfqvist, A., & Gracco, V. L. (2002) Control of oral closure in lingual stop consonant production. *The Journal of the Acoustical Society of America*, *111*(6), 2811-2827.

Lombardi, L. (2002) Coronal epenthesis and markedness. *Phonology*, *19*(2), 219-251.

Ma, J., Chen, X., Wu, Y., & Zhang, L. (2018) Effects of age and sex on voice onset time: Evidence from Mandarin voiceless stops. *Logopedics Phoniatrics Vocology*, 43(2), 56-62.

Macken, M. A., & Barton, D. (1980) The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of child language*, *7*(1), 41-74.

MacKenzie, L., & Tamminga, M. (2021) New and old puzzles in the morphological conditioning of coronal stop deletion. *Language Variation and Change, 33*(2), 217-244. doi:10.1017/S0954394521000119

Maclagan, M. A., Watson, C. I., Harlow, R., King, J., & Keegan, P. 2009. /u/ fronting and /t/ aspiration in Māori and New Zealand English. Language Variation and Change 21:175–192.
Maddieson, I. (1997) Phonetic universals. In J. Laver & W. J. Hardcastle (Eds.), *Handbook of Phonetic Sciences* (pp. 619–639) Blackwells Publishers.

Maddieson, I. (2013) Voicing in Plosives and Fricatives. In M. S. Dryer. & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/4, Accessed on 2022-04-14.)

Mahboob, A. & Ahmar, N.H. (2008) Pakistani English: Phonology. In R. Mesthrie (Ed.) Varieties of English volume 4. Africa, South and Southeast Asia. Berlin & New York: Mouton de Gruyter. pp. 164-176.

Mansfield, J. B. (2015) Consonant lenition as a sociophonetic variable in Murrinh Patha (Australia) *Language Variation and Change*, *27*(2), 203-225.

Mascaró, J. (1989) On the form of segment deletion and insertion rules. *Probus 1*, 31–61.

McCarthy, O., & Stuart-Smith, J. (2013) Ejectives in Scottish English: A social perspective. Journal of the International Phonetic Association, 43(3), 273-298. doi:10.1017/S0025100313000212

McLeod, S., & Crowe, K. (2018) Children's consonant acquisition in 27 languages: A cross-linguistic review. *American journal of speech-language pathology*, *27*(4), 1546-1571.

Mead, J. E. R. E., Turner, J. M., Macklem, P. T. & Little, J. B. (1967) Significance of the relationship between lung recoil and maximum expiratory flow. *Journal of Applied Physiology*, *22*(1), 95-108.

Mees, I. M., & Collins, B. (1999) Cardiff: A real-time study of glottalisation. In Foulkes, P. & Docherty, G. J. (Eds.), Urban voices: Accent studies in the British Isles. London: Arnold. 185–202.

Mesthrie, R. (2008) Indian South African English: phonology. In R. Mesthrie (ed.) Varieties of English volume 4. Africa, South and Southeast Asia. Berlin & New York: Mouton de Gruyter. pp. 188-199.

Mesthrie, R. (2012) Ethnicity, substrate and place: The dynamics of Coloured and Indian English in five South African cities in relation to the variable (t) *Language Variation and Change, 24*(3), 371-395. doi:10.1017/S0954394512000178

Michnowicz, J., & Carpenter, L. (2013) Voiceless stop aspiration in Yucatan Spanish: A sociolinguistic analysis. Spanish in Context, 10(3), 410-437.

Mielke, J., & Nielsen, K. Y. (2018) Voice Onset Time in English voiceless stops is affected by following postvocalic liquids and voiceless onsets. *The Journal of the Acoustical Society of America*, *144*(4), 2166–2177. https://doi.org/10.1121/1.5059493

Miller, J. L., Green, K. P., & Reeves, A. (1986) Speaking rate and segments: A look at the relation between speech production and speech perception for the voicing contrast. *Phonetica* 43 (1–3) : 106. DOI: http://dx.doi.org/10.1159/000261764

Moran, S. & McCloy, D. (Eds.) 2019. PHOIBLE 2.0. Jena: Max Planck Institute for the Science of Human History. (Available online at http://phoible.org, Accessed on 2022-04-22.)

Morris, R. J., & Brown, W. S. (1994) Age-related differences in speech variability among women. *Journal of Communication Disorders*, *27*(1), 49–64. https://doi.org/10.1016/0021-9924(94)90010-8

Morris, R. J., McCrea, C. R., & Herring, K. D. (2008) Voice onset time differences between adult males and females: Isolated syllables. *Journal of Phonetics*, *36*(2), 308–317. https://doi.org/10.1016/j.wocn.2007.06.003

Morris, J., & Hejná, M. (2020) Pre-aspiration in Bethesda Welsh: A sociophonetic analysis. Journal of the International Phonetic Association, 50(2), 168-192.

Mortensen, J., & Tøndering, J. (2013) The effect of vowel height on voice onset time in stop consonants in CV sequences in spontaneous Danish. In R. Eklund (Ed.), *Proceedings of Fonetik* (pp. 49–52).

Nance, C., & Stuart-Smith, J. (2013) Pre-aspiration and post-aspiration in Scottish Gaelic stop consonants. Journal of the International Phonetic Association, 43(2), 129-152. doi:10.1017/S0025100313000042

Nearey, T. M., & Rochet, B. L. (1994) Effects of place of articulation and vowel context on VOT production and perception for French and English stops. *Journal of the International Phonetic Association*, *24*(1), 1–18. https://doi.org/10.1017/S0025100300004965

Neiman, G. S., Klich, R. J., & Shuey, E. M. (1983) Voice onset time in young and 70-year-old women. *Journal of Speech, Language, and Hearing Research*, *26*(1), 118-123.

Nelson, N. R., & Wedel, A. (2017) The phonetic specificity of competition: Contrastive hyperarticulation of voice onset time in conversational English. *Journal of Phonetics*, *64*, 51-70.

Ngom, F. (2003) *Wolof*. Munich: Lincom Europa.

Nielsen, K. (2011) Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39. 132–42. DOI: 10.1016/j.wocn.2010.12.007.

Nodari, R., Celata, C., & Nagy, N. (2019) Socio-indexical phonetic features in the heritage language context: Voiceless stop aspiration in the Calabrian community in Toronto. *Journal of Phonetics*, *73*, 91-112.

O'Dwyer, Fergus (2019) Slit-t in Dublin English. In Villena Ponsoda, J. A., Díaz-Montesinos, F., Ávila-Muñoz, A. M. and Vida-Castro, M. (Eds.) *Language Variation - European Perspectives VII. Selected papers from the Ninth International Conference on Language Variation in Europe (ICLaVE 9)* (pp. 161-175). Amsterdam: John Benjamins.

Ogden, R. A., 2009. An introduction to English phonetics. Edinburgh: Edinburgh University Press.

Oh, E. (2011) Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics* 39 (1) : 59. DOI: http://dx.doi.org/10.1016/j.wocn.2010.11.002

Padgett, Jaye. 2002. Russian voicing assimilation, final devoicing, and the problem of [v]. Ms, UC Santa Cruz. http://people.ucsc.edu/~padgett/papers.html

Paradis, C., & Prunet, J. F. (1991) Introduction: asymmetry and visibility in consonant articulations. In *The special status of coronals: Internal and external evidence* (pp. 1-28) Academic Press.

Parikh, G., & Loizou, P. C. (2005) The influence of noise on vowel and consonant cues. *The Journal of the Acoustical Society of America*, *118*(6), 3874–3888. https://doi.org/10.1121/1.2118407

Parveen, S., & Goberman, A. M. (2014) Presence of stop bursts and multiple bursts in individuals with Parkinson disease. *International Journal of Speech-Language Pathology*, *16*(5), 456–463. https://doi.org/10.3109/17549507.2013.808702

Patrick, P. (2008) Creoles at the intersection of variable processes: -t,d deletion and past-marking in the Jamaican mesolect. *Language Variation and Change, 3*(2), 171-189. doi:10.1017/S095439450000051X

Peng, J. F., Chen, L. M., & Lee, C. C. (2015) Voice onset time of initial stops in Mandarin and Hakka: Effect of gender. *Taiwan Journal of Linguistics*, *28*(9), 63–79.

Percival, M. (2019) Contextual variation in the acoustics of Hul'q'umi'num' ejective stops. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia 2019 (pp. 3270-3274). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Percival, M., Csapó, T. G., Bartók, M., Deme, A., Gráczi, T. E., & Markó, A. (2020) Ultrasound imaging of Hungarian geminates. UltraFest IX.

Percival, M., Kochetov, A., & Kang, Y. (2018) An Ultrasound Study of Gemination in Coronal Stops in Eastern Oromo. In *INTERSPEECH* (pp. 1531-1535).

Petrosino, L., Colcord, R. D., Kurcz, K. B., & Yonker, R. J. (1993) Voice onset time of velar stop productions in aged speakers. *Perceptual and Motor Skills*, *76*(1), 83–88. https://doi.org/10.2466/pms.1993.76.1.83

Pharao, N., & Maegaard, M. (2017) On the influence of coronal sibilants and stops on the perception of social meanings in Copenhagen Danish. *Linguistics*, *55*(5), 1141-1167.

Phillips, A., & Tucker, B. (2020) Context effects on the acoustic realization of stops and affricates in Northern Pwo Karen. Journal of the International Phonetic Association, 1-32. doi:10.1017/S0025100320000109

Piccardi, D. (2017) Sociophonetic factors of speakers' sex differences in Voice Onset Time: A Florentine case study. Fattori sociali e biologici nella variazione fonetica./Social and biological factors in speech variation. Milano: Officinaventuno, 83-106.

Pierrehumbert, J. B. (2006). The next toolkit. *Journal of Phonetics*, *34*(4), 516–530.

Pierrehumbert, J. B., & Talkin, D. (1992) Lenition of /h/ and glottal stop. In G. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 90–117) Cambridge University Press.

Pike, E.V., Forster, K., & Forster, W.J. 1986. Fortis versus lenis consonants in the Paya dialect of Kuna. In Elson, Benjamin F. (ed.), *Language in Global Perspective: Papers in Honor of the 50th Anniversary of the Summer Institute of Linguistics 1935-1985* (pp. 451-464).

Podesva, R., Reynolds, J., Callier, P., & Baptiste, J. (2015) Constraints on the social meaning of released /t/: A production and perception study of U.S. politicians. Language Variation and Change, 27(1), 59-87. doi:10.1017/S0954394514000192

Podesva, R. J., Roberts, S.J., & Campbell-Kibler, K. (2001) Sharing resources and indexing meanings in the production of gay styles. In K. Campbell-Kibler, R. Podesva, S. Roberts & A. Wong (Eds.), *Language and sexuality: Contesting meaning in theory and practice*. Stanford: CSLI Publications. 175–189.

Port, R. F., & Rotunno, R. (1979) Relation between voice-onset time and vowel duration. *The Journal of the Acoustical Society of America*, *66*(3), 654–682. https://doi.org/10.1121/1.383692

Puggaard, R. (2021) Modeling regional variation in voice onset time of Jutlandic varieties of Danish. In H. Van de Velde, N.H. Hilton, R. Knooihuizen (Eds.) Language Variation–European Perspectives VIII: Selected papers from the Tenth International Conference on Language Variation in Europe (ICLaVE 10) John Benjamins Publishing Company. pp. 79-109.

Purse, R. (2019) The articulatory reality of coronal stop "deletion". In S. Calhoun, P. Escudero, M. Tabain & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia 2019 (pp. 1595-1599). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Raymond, W., Brown, E., & Healy, A. (2016) Cumulative context effects and variant lexical representations: Word use and English final t/d deletion. *Language Variation and Change, 28*(2), 175-202. doi:10.1017/S0954394516000041

Recasens, D., & Mira, M. (2015) Place and manner assimilation in Catalan consonant clusters. Journal of the International Phonetic Association, 45(2), 115-147. doi:10.1017/S0025100315000080

Remijsen, B., Ayoker, O. G., & Mills, T. (2011) Shilluk. *Journal of the International Phonetic Association, 41*(1), 111-125.

Riney, T., Takagi, N., Ota, K., & Uchida, Y. (2007) The intermediate degree of VOT in Japanese initial voiceless stops. Journal of Phonetics, 35, 439-444.

Ryalls, J., Zipprer, A., Baldauff, P. (1997) A preliminary investigation of the effects of gender and race on voice onset time. Journal of Speech, Language, and Hearing Research 40, 642-645.

Ryalls, J. H., Simon, M., & Thomason, J. (2004) Voice Onset Time production in older Caucasian- and African-Americans. *Journal of Multilingual Communication Disorders*, *2*(1), 61–67. https://doi.org/10.1080/1476967031000090980

Sanchez, K., Miller, R. M., & Rosenblum, L. D. (2010) Visual influences on alignment to voice onset time. *Journal of Speech, Language, and Hearing Research, 53*(2), 262–272.

Santa Ana, A.O. (1992) Chicano English evidence for the exponential hypothesis: A variable rule pervades lexical phonology. Language Variation and Change 4:275–289.

Schertz, J. (2013) Exaggeration of featural contrasts in clarifications of misheard speech in English. *Journal of Phonetics*, *41*(3-4), 249-263.

Schertz, J., & Johnson, E. K. (2022) Voice onset time imitation in teens versus adults. *Journal of Speech, Language, and Hearing Research*, 1-12.

Schertz, J., & Khan, S. (2020) Acoustic cues in production and perception of the four-way stop laryngeal contrast in Hindi and Urdu. *Journal of Phonetics*, *81*, 100979.

Schleef, E. (2013) Glottal replacement of /t/ in two British capitals: Effects of word frequency and morphological compositionality. *Language Variation and Change, 25*(2), 201-223. doi:10.1017/S0954394513000094

Schwartz, S. R. (2018) *Repping the Streets, Repping the Hometown: A Sociophonetic Analysis of Dialectal Variation in the Moroccan Hip Hop Community*. MA dissertation, University of Texas at Austin. http://hdl.handle.net/2152/63770

Scobbie, J. M. (2006) Flexibility in the face of incompatible English VOT systems In: L. Goldstein, D. H. Whalen, C. T. Best, *Papers in laboratory phonology 8: Varieties of phonological competence*. Berlin: Mouton de Gruyter, pp. 367-392.

Shaw, H. & Foulkes, P. (2015) Real time change in Prince William's speech. Paper presented at UKLVC 10, University of York.

Skarnitzl, R., & Rálišová, D. (2022) Phonetic variation of Irish English /t/ in the syllabic coda. *Journal of the International Phonetic Association*, online first view. doi:10.1017/S0025100321000347

Sleeper, M. (2020) Contact effects on voice-onset time (VOT) in Patagonian Welsh. Journal of the International Phonetic Association, 50(2), 153-167. doi:10.1017/S002510031800021X

Sokolovic-Perovic, M. (2012) *The voicing contrast in Serbian stops.* PhD thesis, Newcastle University.

Stevens, K. N., & Blumstein, S. E. (1981) The Search for Invariant Acoustic Correlates of Phonetic Features. In *Perspectives on the Study of Speech* (pp. 1–38).

Stevens, K. N., & Klatt, D. H. (1974) Role of formant transitions in the voiced-voiceless distinction for stops. *The Journal of the Acoustical Society of America*, 55(3), 653-659.

Stölten, K., & Engstrand, O. 2002. Effects of sex and age in the Arjeplog dialect: A listening

Straw, M., & Patrick, P.L. (2007) Dialect acquisition of glottal variation in /t/: Barbadians in Ipswich. Language Sciences 29:385–407.

Stuart-Smith, J., Sonderegger, M., Rathcke, T. & Macdonald, R. (2015) The private life of stops: VOT in a real-time corpus of spontaneous Glaswegian. Laboratory Phonology, 6(3-4), 505-549. https://doi.org/10.1515/lp-2015-0015

Summerfield, Q. (1975) How a full account of segmental perception depends on prosody and vice versa. In *Structure and process in speech perception* (pp. 51-68) Springer, Berlin, Heidelberg.

Sundara, M. (2005) Acoustic-phonetics of coronal stops: A cross-language study of Canadian English & Canadian French. *The Journal of the Acoustical Society of America*, *118*(2), 1026–1037. https://doi.org/10.1121/1.1953270

Sundara, M., Polka, L., & Baum, S. R. (2006) Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition*, *9*(1), 97–114. https://doi.org/10.1017/S1366728905002403

Suomi, K. (1980) Voicing in English and Finnish stops. Publications of the Department of Finnish and General Linguistics of the University of Turku 10. Ph.D. Dissertation.

Swartz, B. L. (1992) Gender difference in voice onset time. *Perceptual and Motor Skills*, *75*(3), 983–992.

Tagliamonte, S., & Temple, R. A. M. (2005) New perspectives on an ol' variable: (t, d) in British English. *Language Variation and Change*, *17*(3), 281–302.

Temple, R. (1999) Phonetic and sociophonetic conditioning of voicing patterns in the stop consonants of French. In *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 1409-1412).

Themistocleous, C. (2016) The bursts of stops can convey dialectal information. The Journal of the Acoustical Society of America, 140(4), EL334-EL339.

Theodore, R. M., & Miller, J. L. (2010) Characteristics of listener sensitivity to talker-specific phonetic detail. *The Journal of the Acoustical Society of America*, *128*(4), 2090-2099.

Theodore, R. M., Miller, J. L., & DeSteno, D. (2009) Individual talker differences in voice-onset-time: Contextual influences. *The Journal of the Acoustical Society of America*, *125*(6), 3974. https://doi.org/10.1121/1.3106131

Torre, P., Barlow J.A. 2009. Age-related changes in acoustic characteristics of adult speech. Journal of Communication Disorders 42(5), 324-333.

Toscano, J. C., & McMurray, B. (2010) Cue integration with categories: Weighting acoustic cues in speech using unsupervised learning and distributional statistics. *Cognitive science*, *34*(3), 434-464.

Toscano, J. C., McMurray, B., Dennhardt, J., & Luck, S. J. (2010) Continuous perception and graded categorization: Electrophysiological evidence for a linear relationship between the acoustic signal and perceptual encoding of speech. *Psychological Science*, *21*(10), 1532-1540.

Trumper, J., & Maddalon, M. (1990) Il problema delle varietà: L'italiano parlato nel Veneto. In M. A. Cortelazzo, & A. M. Mioni (Eds.), L'italiano regionale. Atti del XVIII congresso internazionale di studi della Società di Linguistica Italiana (pp. 159-191) Roma: Bulzoni.

Utsugi, A., Sasaki, K., & Igarashi, Y. (2013) Regional variation of VOT in Ibaraki Japanese. In *Proceedings of the 27th annual meeting of the Phonetic Society of Japan* (pp. 119-124) 日本音声学会.

van Alphen, P. M., & Smits, R. (2004) Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics*, *32*(4), 455–491. https://doi.org/10.1016/j.wocn.2004.05.001

van Rooy, B. (2008) Black South African English: Phonology. In R. Mesthrie (ed.) Varieties of English, volume 4. Africa, South and Southeast Asia. Berlin & New York: Mouton de Gruyter. pp. 177-187.

Vicenik, C. (2008) *An Acoustic Analysis of Georgian Stops*. Master's Thesis. UCLA.

Vicenik, C. (2010) An acoustic study of Georgian stop consonants. *Journal of the International Phonetic Association*, 40(1), 59-92. doi:10.1017/S0025100309990302

Vietti, A. (2019) Phonological variation and change in Italian. In *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press.

Vihman, M.M. (1996) *Phonological Development: The Origins of Language in the Child*. Oxford: Blackwell.

Villafaña Dalcher, C. (2008) Consonant weakening in Florentine Italian: a cross-disciplinary approach to gradient and variable sound change. Language Variation and Change 20: 275–316.

Volenec, V., & Reiss, C. (2017) Cognitive phonetics: The transduction of distinctive features at the phonology-phonetics interface. *Biolinguistics*, *11.SI*, 251–294.

Wade, L., Lai, W., & Tamminga, M. (2021) The reliability of individual differences in VOT imitation. *Language and Speech*, *64*(3), 576-593.

Wang, X. (2019) Dialect contact across three generations: a sociophonetic analysis of variation in [p^h t^h k^h, h] in a contact variety in Hohhot, China. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia (pp. 2339-2343) Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Warner, N., & Weber, A. (2001) Perception of epenthetic stops. *Journal of Phonetics*, *29*, 53-87.

Warner, N., & Weber, A. (2002) Stop epenthesis at syllable boundaries. In *7th International Conference on Spoken Language Processing [ICSLP2002-INTERSPEECH 2002]* (pp. 1121-1124) ISCA Archive.

Watson, J. C. (2002) The phonology and morphology of Arabic. OUP Oxford.

Weismer, G. (1979) Sensitivity of voice-onset time (VOT) measures to certain segmental features in speech production. *Journal of Phonetics*, *7*, 197–204.

Wells, John C. (1982) *Accents of English* (3 vols.) Cambridge: Cambridge University Press.

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

Whisker-Taylor, K., & Clark, L. (2019) Yorkshire Assimilation: Exploring the Production and Perception of a Geographically Restricted Variable. *Journal of English Linguistics*, *47*(3), 221-248.

Whiteside, S. P., Henry, L., & Dobbin, R. (2004) Sex differences in voice onset time: A developmental study of phonetic context effects in British English. The Journal of the Acoustical Society of America, 116(2), 1179-1183.

Whiteside, S. P., & Marshall, J. (2001) Developmental trends in voice onset time: Some evidence for sex differences. *Phonetica*, *58*(3), 196–210. https://doi.org/10.1159/000056199

Wiese, R. (2000) *The Phonology of German*. Oxford: Oxford University Press.

Yağlı, E. (2018) Indexing social meaning: Sociophonetic variables and listener perceptions of Turkish. PhD thesis, Hacettepe University.

Yang, J. (2018) Development of stop consonants in three-to six-year-old Mandarin-speaking children. *Journal of Child Language*, *45*(5), 1091–1115. https://doi.org/10.1017/S0305000918000090

Yao, Y. (2007) Closure duration and VOT of word-initial voiceless plosives in English in spontaneous connected speech. *UC Berkeley Phonology Lab Annual Report*, 183–225.

Yao, Y. (2009) Understanding VOT variation in spontaneous speech. In M. Pak (Ed.), *Current Numbers in Unity and Diversity of Languages* (pp. 1122–1137) Linguistic Society of Korea.

Yoo, I., & Blankenship, B. (2003) Duration of epenthetic [t] in polysyllabic American English words. *Journal of the International Phonetic Association, 33*(2), 153-164. doi:10.1017/S0025100303001269

Yu, A. C., Abrego-Collier, C., & Sonderegger, M. (2013) Phonetic imitation from an individual-difference perspective: Subjective attitude, personality and "autistic" traits. *PloS one*, *8*(9), e74746.

Zerbe, A. (2013) P'urépecha fortis v. lenis consonants. *Working Papers in Linguistics (UWWPL)*, *31.* University of Washington.

Zlatin, M. A., & Koenigsknecht, R. A. (1976) Development of the voicing contrast: A comparison of voice onset time in stop perception and production. *Journal of Speech and Hearing Research*, *19*(1), 93–111.

Zsiga, E.C. (2011) Local assimilation. In M. van Oostendorp, C. Ewen, B. Hume, and K. Rice (Eds.) The Companion to Phonology, Volume 3. Wiley-Blackwell.

Zue, V. W. (1976) *Acoustic Characteristics of Stop Consonants: A Controlled Study*. PhD thesis, Massachusetts Institute of Technology.