Focus on Consonants: Prosodic Prominence and the Fortis-Lenis Contrast in English

Míša Hejná & Anna Jespersen
Aarhus University

Abstract
This study investigates the effects of intonational focus on the implementation of the fortis-lenis contrast. We analyse data from 5 speakers of different English dialects (Ocke's colleagues), with the aim of examining the extent to which different correlates of the contrast are used by each speaker, and whether the contrast is implemented differently across different levels of focal prominence (narrow focus, broad focus, de-accentuation). The correlates examined include three measures often associated with the contrast (pre-obstruent vowel duration, consonant/vowel durational ratio, rate of application of obstruent voicing), as well as a number of lesser-investigated phenomena. Firstly, we find that individual speakers utilise different phonetic correlates to implement the fortis-lenis contrast. Secondly, focus affects several of these, with the biggest effect found with consonant/vowel ratio, and the smallest with obstruent voicing.

1. Introduction

It has been frequently claimed that there is more variation in vowels than consonants (e.g. Bohn & Caudery, 2017, p. 63), possibly because "consonantal variation (in British English at least) tends to be used less as a way of marking local identity than vocalic variation does" (Trousdale, 2010, p. 116). An alternative claim may be that "[c]onsonantal features have been studied far less rigorously than vowel features" (Cox & Palethorpe, 2007, p. 342, who comment on the state of consonantal variation studies in Australian English; but see also Su, 2007, p. 6).

Anne Mette Nyvad, Michaela Hejná, Anders Højen, Anna Bothe Jespersen & Mette Hjortshøj Sørensen (Eds.), ***A Sound Approach to Language Matters – In Honor of Ocke-Schwen Bohn*** (pp. 237-270). Dept. of English, School of Communication & Culture, Aarhus University.

These claims are nevertheless somewhat surprising at least in the context of the fortis-lenis contrast, which has traditionally received widespread attention from phoneticians and which has been reported to show a wide range of phonetic implementation strategies (e.g. Jansen, 2004; Kingston, Diehl, Kirk, & Castleman, 2008; Kohler, 1984; Toscano & McMurray, 2010). Limiting ourselves to English, this contrast is generally reported to be cued and signalled by differences in the voicing of the obstruent, the duration of the preceding vowel, and by the presence and duration of post-aspiration (Bohn & Caudery, 2017, pp. 74-81). Regarding plosives, previous work has tended to focus on post-aspiration in the form of VOT. VOT analyses have targeted primarily the word-initial position (as in ***t****ock* vs ***d****ock*; see e.g. discussions in Docherty, 1992; Iverson & Salmons, 2006). For obstruents more generally, analyses have centred on the preceding vowel duration non-initially (e.g. Fox & Palethorpe, 2007, p. 343; Wells, 1990). In comparison to VOT and preceding vowel duration, discussions of *voicing* tend to be limited (but see e.g. Scobbie, 2005; Smith, 1997). More recently, glottalisation and pre-aspiration have been reported as correlates of the contrast in some British English accents (Hejná, 2016b; Hejná & Kimper, In press; Gordeeva & Scobbie, 2013), and glottalisation also in Australian English (Penney, Cox, Miles & Palethorpe 2018). Furthermore, a range of other potential correlates of the fortis-lenis contrast in British English has been put forward (Hejná & Kimper, Submitted). Thus the implementation strategies of the fortis-lenis contrast are abundant in English varieties.

We propose that one way to tap into the strength of different correlates of the fortis-lenis distinction is through investigating its variation under different levels of prosodic prominence. Previous work on the fortis-lenis contrast has tended to make use of laboratory speech, in which the target word for the investigation is often prosodically focused (see e.g. the following studies, which use word lists and/or carrier sentences likely to encourage production with narrow focus and which represent but a few examples available: Gordeeva & Scobbie, 2013; Hejná, 2016b; Hejná & Kimper, In press; Scobbie, 2005; Smith, 1997). Apart from the fact that these conditions may not reflect the phonetic situation in more natural data regarding VOT, voicing, and the preceding vowel duration, it is also possible that the robustness of lesser-investigated potential correlates in such studies has a causative relationship with prosodic prominence (as discussed e.g. in Mücke & Grice, 2014). Controlling for this variable may thus yield insights into the degree to which individual correlates

of the fortis-lenis contrast are utilised by speakers of English. At the same time, such an investigation can enrich our knowledge of prosodic prominence. More specifically, non-focal prominence, such as word stress, is known to have different acoustic correlates across languages (Mennen, 2006; Leemann, Kolly, Li, Chan, Kwen, & Jespersen, 2015). While the contributions of syllable duration is known to play a role in the perception of prominence (Leemann et al., 2015; Hua, Li, & Wayland, this volume), little is known of the contribution of segmental variation to prosodic prominence in general, and to intonational focus in particular. As such, focusing on focused consonants therefore also extends our knowledge of the ways in which prosodic prominence depends on segmental variation.

1.1. Interactions between prosody and segments

Studies and models of intonational variation, such as variation in intonational focus, tend to treat its physical manifestation as more or less synonymous with variation in f0 (Fundamental Frequency; see Niebuhr, 2013, for a discussion of this). However, intonation is crucially dependable on variation not just in f0 but also in the duration, intensity and phonatory characteristics of segments which form the underlay of suprasegmental variation (Nolan, 2006, p. 433). It has long been known that the intrinsic prosodic characteristics of individual segments can influence the course of f0 movements (Kingston, 1986; Kingston, Diehl, Kirk, & Castleman, 2008). Nonetheless, recent work has started to broaden our knowledge of the ways in which both the production and the perception of prosodic phenomena such as intonation rely on segmental variation. For instance, evidence from German suggests that local intonational categories are cued through variation in the levels of intensity and duration of accented syllables, and that listeners are able to make use of this variation to identify distinct pitch contours and decode their communicative content (Niebuhr & Pfitzinger, 2010; see also Niebuhr, this volume).

There is, however, limited evidence for the impact of intonational variation on segmental realisation. Niebuhr and colleagues have conducted a number of studies on German in which voiceless fricatives such as [f], [s], [ʃ], and [x] have been shown to have greater intensity and higher centres of gravity after rising f0 than after f0 falls, both medially and finally (Niebuhr, 2012; 2013, p. 10). Coarticulatory processes can also be affected: the assimilation of /s/ to /ʃ/ has been argued to be stronger with f0 rises than with falls (Niebuhr et al., 2011). Studies have also suggested

that listeners are able to tap into segmental variation as additional cues to intonational contrasts and to compensate for truncated f0 movements (Kohler, 2011). Limited work has nevertheless been carried out on non-fricative obstruents. One exception to this is Niebuhr (2008), who shows that the aspiration of utterance-final /t/ can be impacted by intonational contour type. Two falling contours were included, which were distinct in the alignment of the peak: early (H+L* L%) or late (L*+H L%). With early f0 peaks, /t/ aspirations were realised as short, intense and with low-frequency centres of gravity, while late-peak aspirations were longer, less intense and featured energy at higher frequencies. A subsequent perception study indicated that listeners made use of /t/ aspiration as a cue to the intended f0 contour.

This paper aims to extend previous findings on the effects of intonational variation on the realisation of consonants by homing in on the contribution of *prosodic focus* on obstruent realisation. The manifestation of focus can have three different outcomes on individual syllables. *Broad focus* assigns the whole utterance prominence.[1] In other words, no single element of the utterance receives special prominence; a nuclear intonation contour is thus assigned to the last metrically prominent syllable in the utterance. This level of focus is perceived by listeners as prosodically neutral (Cruttenden, 1996, p. 87). *Narrow focus* assigns the nucleus to any syllable in the utterance. A syllable receiving narrow focus is highlighted as carrying new information and thus receives a greater amount of prosodic prominence than the nuclear syllable under broad focus (though in phonetic terms, this distinction might not always be unambiguous; cf. Ladd, 1996, pp. 254-56). Assigning narrow focus to a syllable affects the prosodic makeup of the entire utterance, in that it de-accents all following pitch accents (Ladd, 1996, p. 225). *De-accenting* affects any material following a narrowly focussed syllable by reducing e.g. the pitch range of f0 movements and the duration of metrically prominent syllables, and may also affect the strength of boundary cues (Baltazani & Jun, 1999; Norcliffe & Jaeger, 2005). Realisational effects of de-accenting a syllable include reduced vowel durations, changes in vowel quality, and the non-elision of linking- and intrusive /r/ (Ladd, 1996, pp. 266-67). Yet studies have not focused on the fine-grained

[1] Note that "prominence" is used in this study to mean *metrical* prominence; whether and to which extent accented syllables always receive *physical* prominence, as expressed through e.g. f0, durational and segment-realisational means, is not fully established. For a discussion definitions of prominence, see Wagner et al. (2015).

phonetic consequences of the deaccentuation of syllables on consonants, or compared consonantal realisations across different focus conditions (for an overview of papers discussing the effects of focal prominence on vowel realisation, see Mücke & Grice, 2014, p. 5).

1.2. Hypotheses and research questions

This paper aims to answer three research questions:

- Is the implementation of the fortis-lenis obstruent contrast affected by focus?
- In which other ways does focus affect variation in obstruent realisation?
- Does the nuclear intonation contour affect variation in obstruents?

With regard to the first research question, we hypothesise that focus will affect the implementation of the fortis-lenis contrast in such a way that increased levels of prosodic prominence will correlate with higher strength of the consonantal correlates,[2] with a decrease in this strength as the prosodic prominence decreases (Kügler, 2008; Görs & Niebuhr, 2012). This hypothesis is based on the well-known durational differences between different levels of focal prominence (e.g. between broad and narrow focus, see Baumann, Becker, Grice, & Mücke, 2007), and on studies which find a correlation between increased focal prominence and exaggerated articulation, in which a narrowed focus domain is signalled through increased articulatory effort, e.g. through increased lip-rounding in the production of French vowels (Dohen, Loevenbruck, & Hill, 2006; de Jong, 1995) and increased peripherality in German vowels (Baumann et al., 2007).

As to the second research question, it is well-known that there is more variation in obstruent variation in non-foot-initial position (e.g. Smith, 2002), but little is known about other prosodic contexts. Based on such findings, we expect that increased levels of prosodic prominence will correlate with increasing differentiation between the fortis and lenis obstruent realisation. A related hypothesis has previously been suggested for pre-aspiration by Hejná (2015, pp. 241-2), who proposes that pre-aspiration first innovates in lexically stressed rather than unstressed

[2] We follow MacWhinney's definition of correlate strength, which includes correlate/cue reliability and correlate/cue availability (2001, 2012).

syllables. The motivations for this may be related to the durational properties of the preceding vowel, whose quality may be less likely to be affected by pre-aspiration-induced breathiness when they are durationally long (Steriade, 1998, p. 214). If that is the case, pre-aspiration should be likely to interact with other types of prominence, such as focus. To the best of our knowledge, our study is the first to investigate this possibility.

Finally, we will briefly investigate the effects of different nuclear contours on obstruent realisation. This investigation was motivated by findings reported for German by Niebuhr et al. as discussed above. Hypotheses based on this work include a tendency for increased differentiation of the fortis-lenis contrast with high or rising intonational movements, as well as correlation between the complexity and duration of intonational contours. This has often been referred to in the literature (e.g. Ohala, 1978), but not investigated experimentally.

The structure of the study is as follows. We present our methodology for the study, where we also introduce the consonantal variables under consideration. Next, we show our results, examining both individual-specific patterns as well as those shared by our five speakers, and we discuss their implications for the study of the fortis-lenis contrast in English.

2. Methodology

2.1. Data

The data were recorded using a H4 Zoom Handy recorder with a C520 AKG headset microphone at a 44.1 kHz rate with a resolution of 32 bits. We collected a list of words embedded in carrier sentences with different prominence conditions (see below on the specifics of these carrier sentences). These words targeted foot-medial and foot-final fortis and lenis obstruents, namely the alveolar plosive and fricative fortis-lenis pairs /t/-/d/ and /s/-/z/, as shown in Table 1. Thus, each participant produced 24 word types, each repeated six times (ideally three times under narrow focus and three times in a de-accented prosodic environment). Efforts were made to include words with comparable lexical frequencies; however, in this study we were primarily concerned about excluding obviously low frequency items. We obtained 756 tokens for analyses in total.

/s/	/z/	/t/	/d/
bus	*buzz*	*mutt*	*mud*
moss	*Oz*	*lot*	*odd*
lass	*jazz*	*mat*	*mad*
buses	*buzzer*	*mutter*	*muddy*
mossy	*Ozzy*	*otter*	*odder*
lassie	*jazzy*	*matter*	*madder*

Table 1. Words with fortis-lenis alveolar obstruents.

The participants also read a number of distractors, which included *abbey*, *bleak*, *blob*, *blobby*, *goofy*, *Hobbit*, *hobby*, *hoof*, *lab*, *leak*, and *leek*. The order of the items was randomised, which also extends to the different carrier sentences. However, the order in which the words were presented to the participants was uniform across these participants.

We investigated the effects of focus on segmental realisation by incorporating target words into carrier sentences such as *Ocke is writing a paper on the word _?*. In contrast to other studies of focus (e.g. Baltazani & Jun, 1999; D'Imperio, 2001; Xu & Xu, 2005), this experiment thus contained target words appearing at the end of the carrier sentence. That is, the location of the target word in the sentence was not changed across elicitations, and was not generally sentence-medial. We chose this context for two reasons: firstly, it allowed for an investigation of the effect of different nuclear intonation contours on the segmental material. Secondly, the phrase-final context made for a more conservative study. This is because English is known to have significant lengthening of both vowels and non-plosive consonants in phrase-final words (Klatt, 1975; Turk & Shattuck-Hufnagel, 2007). Differences in duration resulting from alterations of focal prominence will therefore be smaller in phrase-final than non-phrase-final positions, all other things being equal. As such, any durational-related differences in consonantal realisation resulting from different focus conditions in this position can therefore be considered robust.

We originally aimed to elicit two extremes in terms of focus: narrow focus, where the word under investigation is made metrically prominent, and de-accentuation, where the target word is made metrically non-prominent. The two focus levels were elicited by 1. capitalising

the target word (*Ocke is writing a paper on the word MOSS?*, where *moss* is the target), or 2. capitalising a non-target word in the carrier sentence (*OCKE is writing a paper on the word moss?*, where *moss* is the target). In order to avoid target words being judged by speakers as new information due to their continued change in static carrier phrases, and thus triggering narrow focus, the two content words in the carrier sentence (in the above sentences, *Ocke* and *paper*) were randomly varied with each new sentence, so that the subject of the sentence was either a pronoun or the name of one of Ocke Bohn's colleagues at the Department of English, and the object a type of research output (*paper*, *keynote*) which could plausibly be generated by these persons[3]. In order to further prompt the speakers to produce the intended level of focus, they were presented with powerpoint slides which contained cues to the appropriate pragmatic context (e.g. "not *Anna* / *Míša*?" or "not the word *wiggle* / *bumblebee* / *pumpkin*?"). Finally, the speakers were given a training set of example sentences prior to the start of the recordings in which the relevant pragmatic contexts were explained to them by the first author.

Despite the written cues and oral demonstrations, when presented with the carrier phrases intended to elicit narrow focus, several speakers varied between producing the sentences with narrow focus on the target word and with broad (i.e. "neutral") focus. In this case, the target word, because of its placement at the end of the target sentence in this experiment, corresponded to the nuclear accent. Broad focus on the sentence in this case meant that the target word became more prominent than in the deaccented condition, but did not receive extra levels of metrical prominence due to narrow focus. In this way, we inadvertently ended up eliciting three levels of focus, whereby the target words were either de-accented, relatively neutral (but accented) or under narrow intonational focus.

2.2. Quantifying consonantal variation

In order to answer the research questions, we measured the duration of the pre-obstruent vowel, the duration of the obstruent, and the frequency of application of voicing. In addition, the following phenomena were annotated for their presence: post-aspiration, affrication, spirantisation, ejectivisation, glottal replacement, pre-aspiration, and flapping. All annotation was carried out manually in Praat (Boersma & Weenink,

[3] See Niebuhr & Michaud (2015) for further discussion of read speech and intonational elicitation.

1992-2017). Since none of the speakers displayed ejectivisation and glottal replacement, these phenomena are not commented on in what follows. Furthermore, glottal reinforcement presented challenges which were beyond the scope of this paper, and was therefore not quantified.

Segmental annotation

On the whole, the onset and offset of the vowel were defined as the onset and offset of periodicity, which were determined on the basis of the soundwave. However, further context-specific criteria were used. Firstly, word-initial affricates (*jazz*, *jazzy*) are generally voiceless; however, they could be articulated as voiced at their offset, which means that a period of voiced oral friction could be found in our affricate-vowel sequences. In such cases, it was the offset of the oral friction that defined the onset of the vowel. Secondly, when a vowel was followed by a phonetically voiced fricative (e.g. *Oz*, *Ozzie*), it was again the onset of the oral friction of that fricative that defined the offset of the vowel. Next, in case of vowel-initial words, aperiodic glottalisation was excluded from the vocalic interval (*otter*, *odder*, *odd*, *Oz*, *Ozzie*), unless this glottalisation actually affected the whole vowel.[4] On the other hand, where glottalisation was found which was not due to vowel-initial word effects, this was always considered part of the vowel for practical rather than theoretical reasons: the boundary has to be placed somewhere in order to extract the measures and it is not always obvious whether the glottalisation is conditioned by the following obstruent.[5]

Voicing

We identified the presence and the duration of voicing on the basis of the waveform. The onset of periodicity was considered the onset of voicing and its offset was considered the offset of voicing. As long as there was some voicing present in the obstruent (fortis or lenis), the token was

[4] This was done because it became obvious that glottalisation could reach fairly high durational values without obviously affecting those of the vowel. The entire vowel was sometimes glottalised, in which case glottalisation could not be excluded as this would have left us with no vowel to measure. Yet these cases were *very* infrequent.

[5] Although pre-glottalisation is a frequent feature of fortis plosives in a number of English accents and has, moreover, been found to function as a correlate of the fortis-lenis contrast in at least three accents of British English (see Hejná & Kimper, In press), this phenomenon was excluded from our analyses. This decision was made because it is problematic to distinguish utterance-final glottalisation, individual-specific global glottalisation, and subsegmental glottalisation (i.e. glottal reinforcement).

considered to contain voicing. Although we do not report on variation as to the extent to which the obstruent is voiced (e.g. 50% of the obstruent), the vast majority of the phonetically voiced obstruents are only partially voiced rather than fully voiced. As can be expected, there is variation in the exact proportion value across and within our speakers. We also find a considerable amount of variation as to where in the consonantal interval voicing occurs (only initially, only finally, both initially and finally, throughout the duration of the obstruent). However, due to the limitations of space, variation beyond presence/absence of voicing is not commented on further in this study.

Pre-aspiration

Pre-aspiration is defined as a period of voiceless (primarily) glottal friction which occurs in sequences of a vowel and a phonetically voiceless obstruent, in line with Hejná (2015; 2016a, amongst others), as shown in Figure 1. It is therefore distinguished from local breathiness, which is very closely linked to voiceless pre-aspiration (see Hejná, 2016a, for more details).[6] We followed the same segmentational criteria as Hejná (2016b), who looked into the role of pre-aspiration in the fortis-lenis plosive contrast in Aberystwyth English. Regarding the segmentation criteria, see Hejná (2015, pp. 85-87). Post-aspiration is defined and identified in the same way as pre-aspiration, with the difference of post-aspiration occurring during the release phase of the obstruent rather than prior to it.

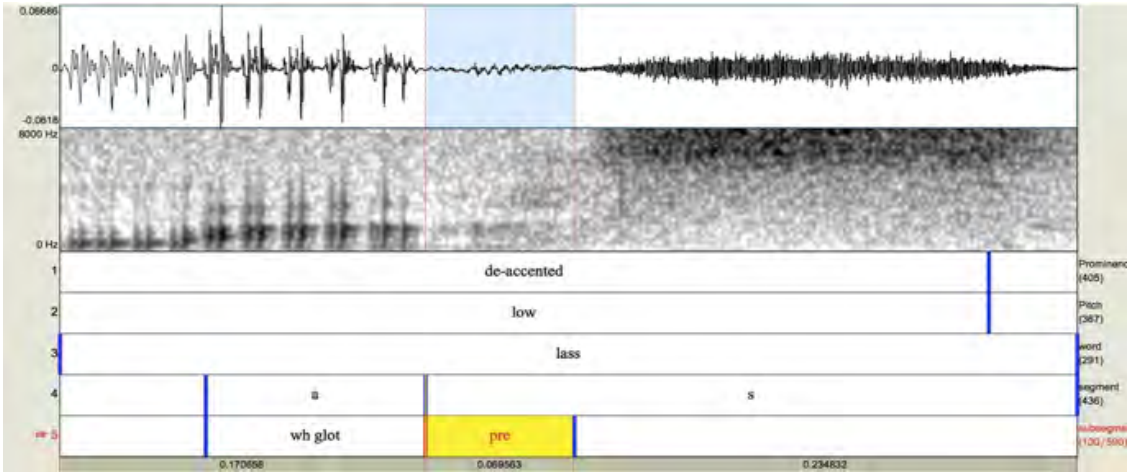


Figure 1. Segmentation of pre-aspiration in the fricative context (illustrated with Elly).

[6] Due to various aspects of breathiness in our data which make the phenomenon fairly complex this paper excludes discussions of breathiness.

Flapping/tapping

Following Bohn and Caudery (2017, p. 79), and unlike Ladefoged (1968), we do not distinguish flapping and tapping. The term flapping is used in the rest of the paper. The phenomenon was identified on the basis of absence of the plosive burst and the presence of voicing, as illustrated in Figure 2 below.

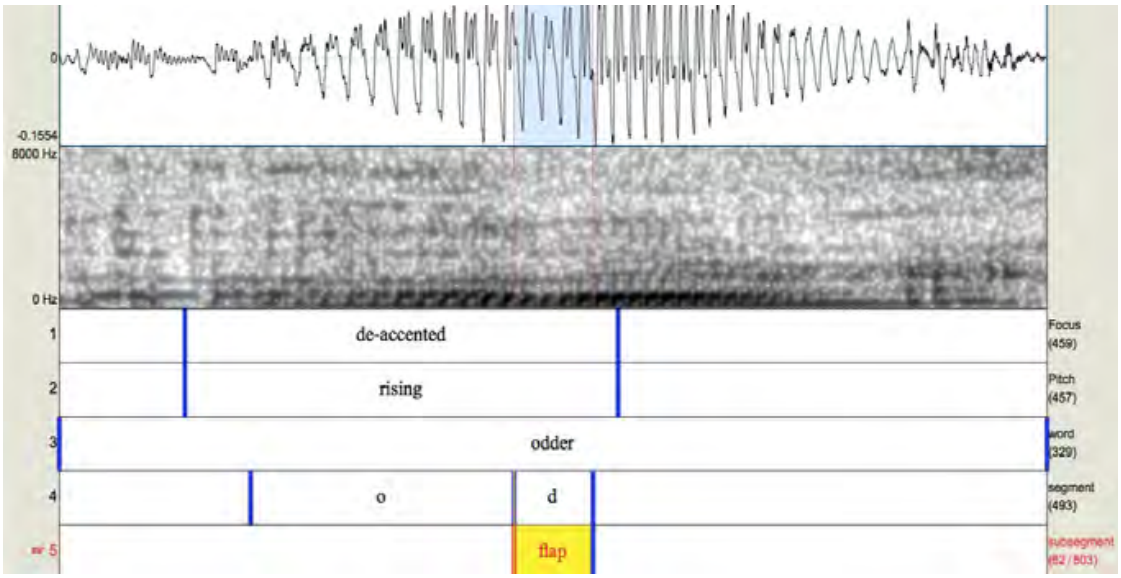


Figure 2. Identification of flapping (illustrated with Mark).

The vast majority of the potential cases of flaps at hand clearly met these criteria and were therefore unambiguously labelled as flaps. Similarly, the vast majority of the phonological plosives met the criteria used to identify voiced plosives. Nevertheless, it is noteworthy (but not surprising) that there is variation in the overall duration of voiced plosives in the data, resulting in what seems to be a durational continuum between a flap and a voiced plosive, where it can be solely the presence of a burst that distinguishes a flap from a plosive. This is found not only across the individual speakers but also within the individuals. For the purposes of this study, as long as a burst was clearly identifiable, the obstruent was classified as a plosive rather than a flap.

Affrication

Affrication was identified by the presence of higher intensity energy in higher frequencies, which indicates oral rather than glottal friction (Figure 3). Affricated plosives were occasionally post-aspirated as well.

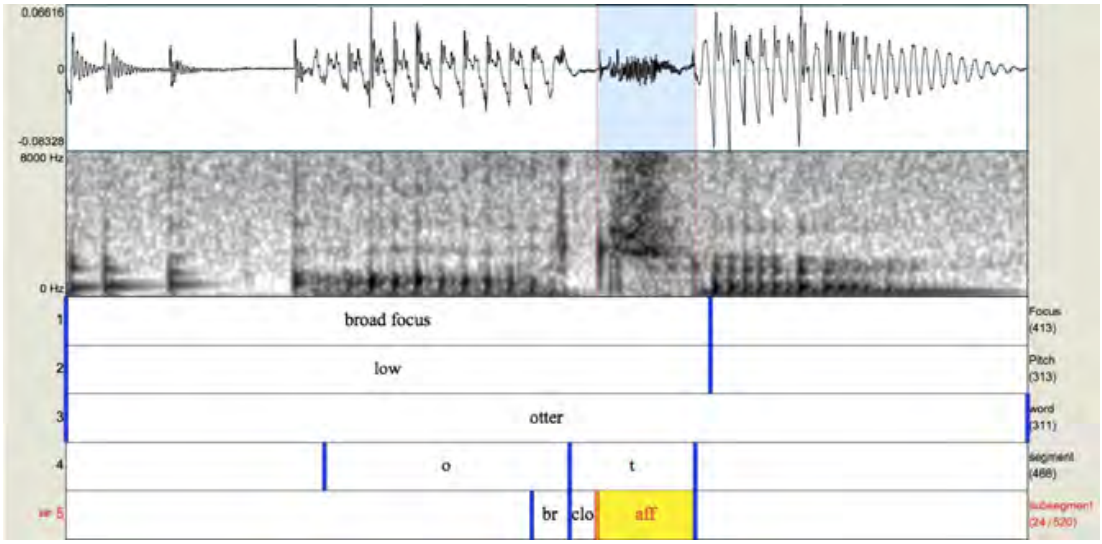


Figure 3. Identification of affrication (illustrated with Stephen).

Spirantisation

Spirantisation of plosives was identified on the basis of the absence of the burst of the plosive. This could result either in a semi-spirantised plosive (semi-fricative; see Stevens & Hajek, 2005) or in a fully spirantised plosive, i.e. in a fricative (Figure 4 for full spirantisation). Interestingly, we found cases of fully spirantised plosives which were also pre- and/or post-aspirated.

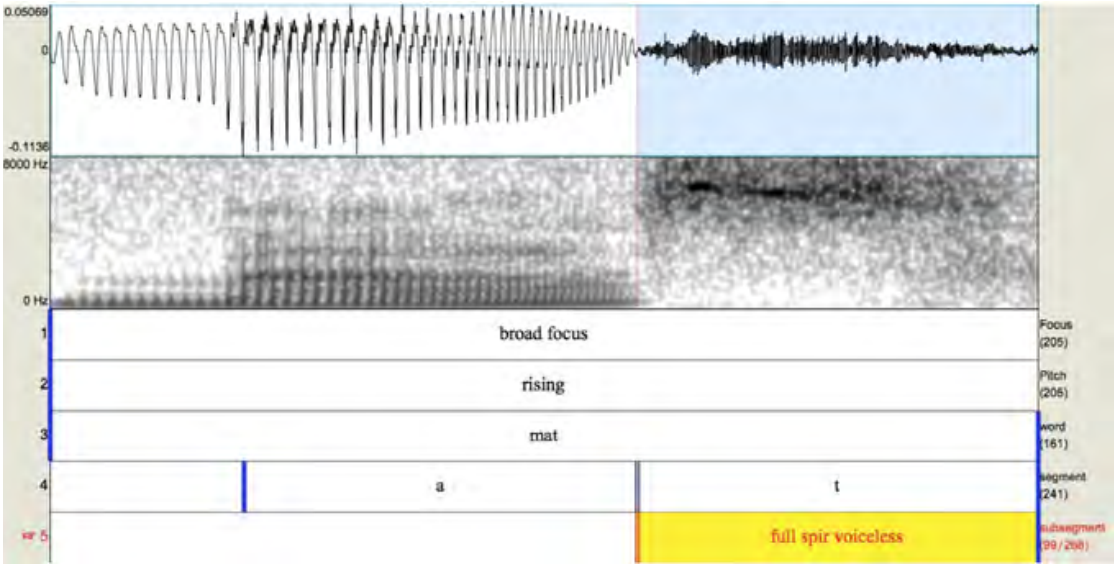


Figure 4. Identification of full spirantisation (illustrated with Antoinette).

2.3. Speakers

5 native speakers of English were recorded reading the words in the appropriate carrier sentences. These speakers were Ocke Bohn's colleagues, all of whom worked at the Department of English, Aarhus University, at the time of the recording (spring 2018). The five bravehearts include Antoinette Fage-Butler, Elly McCausland, Mark Eaton, Stephen Joyce, and Sophia Kier-Byfield. Since the speakers were born and raised in different regions, and have had variable life histories due to their profession as academics, their dialectological background is not uniform. Antoinette and Stephen represent Irish English (1 female speaker, 1 male speaker); Elly represents Standard Southern British English (1 female speaker); Mark represents Canadian English (1 male speaker); and Sophia represents Welsh English (1 female speaker), although Sophia's background and dialect history are more complex than that of the other participants.[7] Apart from regional and sex differences, the five departmental members also differ in other dimensions, most notably that of age. Whilst we acknowledge that this most likely accounts for much of the variation found in the dataset, these individual differences do not present a problem for our research questions.

2.4. Data processing

Statistical analyses were done through linear and logistic mixed effects and random forest modelling in R (R Core Team, 2018) and RStudio (RStudio Team, 2015), using the packages *lmer* (Kuznetsova et al., 2017), *lme4* (Bates et al., 2015), *effects* (Fox, 2003), *ranger* (Wright & Ziegler, 2017), and *lattice* (Sakar, 2008). Regression models were compared hierarchically using the *aov()* function. Post-hoc Tukey tests were performed through the *TukeyHSD()* function.

3. Results

In this section we give broad overviews of the main correlates of the fortis-lenis contrast as found across the speakers, but, given the highly variable speaker sample, also closely inspect potential individual patterns. We then examine the ways in which prosodic variation affects these realisations. We consider other correlates of the contrast, and then investigate the effects of focus and contour type on these lesser-investigated correlates as produced by five members of the AU Department of English.

---

[7] She spent most of her childhood in Wales and her parents are not speakers of a British English variety.

3.1. Preceding vowel duration, vowel-consonant ratio, and voicing of the obstruent

Firstly, the preceding vowel duration distinguishes the fortis-lenis obstruent contrast in all five department members (Figure 5): the duration of the pre-obstruent vowel is longer in the lenis than in the fortis obstruents, although this difference is not significant ($F(3, 23)=1.21$, $SS=2993.9$, $p=.33$).[8] /t/ is associated with the shortest preceding vowel duration in our sample ($\overline{x}$=129ms, SD=40.9ms), as compared to /d/ ($\overline{x}$=155.7ms, SD=60ms) and /s/ ($\overline{x}$=139.5ms, SD=42.9 ms).

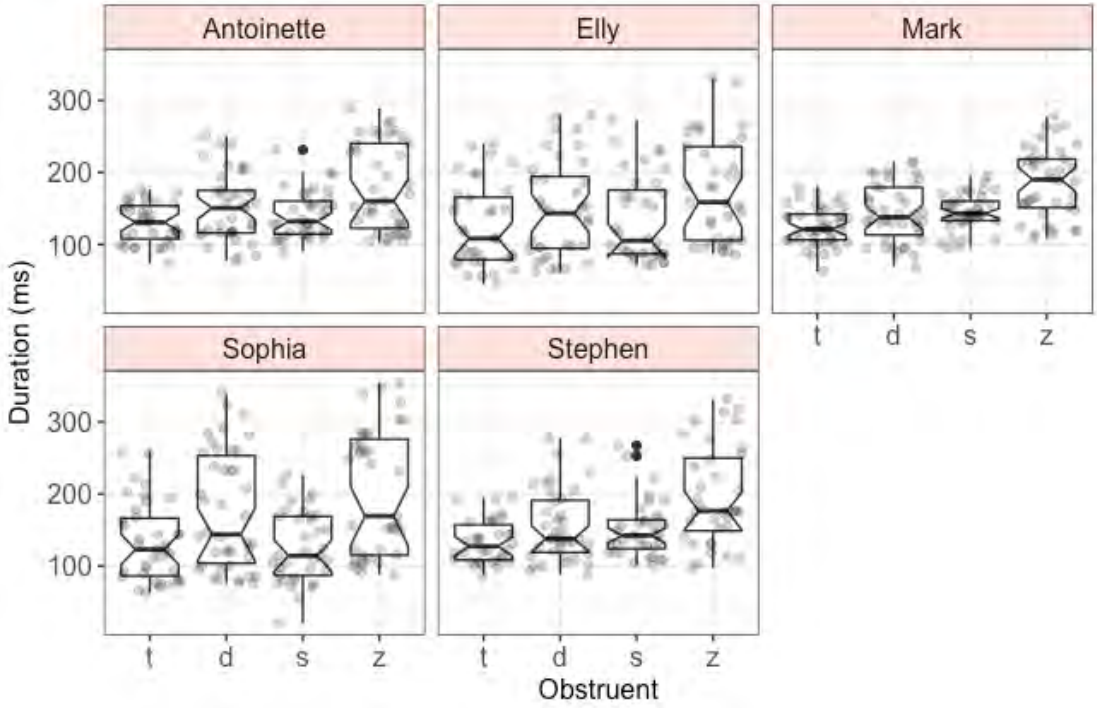


Figure 5. Preceding vowel duration (ms) and the fortis-lenis contrast by individual.

Tukey post-hoc tests revealed that the only statistically significant difference between individual obstruents is between /t/ and /z/ ($p<.01$). Furthermore, we also see individual variation in the magnitude of these differences.[9] As shown in Figure 6, when we consider the V/C ratio and its role in the fortis-lenis contrast implementation, the duration of the preceding vowel is always longer than that of the obstruent in the lenis

[8] Final model: lmer(Vdur ~ consonant + (1|speaker) + (1|word), data=data).

[9] This individual variation is not due to potential differences in speaking rate: the same results are obtained with normalised vowel durations (as a percentage of the overall word duration).

series: /t/ ($\bar{x}$=1.31ms, *SD*=0.94ms) versus /d/ ($\bar{x}$=2.43ms, *SD*=1.1ms), and /s/ ($\bar{x}$=0.84ms, *SD*=0.24ms) versus /z/ ($\bar{x}$=1.6ms, *SD*=0.47ms). Consonant as a factor is thus a highly significant predictor of V/C ratio ($F(3, 23)=20.457$, *SS*=24.54, $p<0.0001$)[10], and Tukey tests show that the fortis/lenis members in each pair are significantly different from each other (/t/-/d/: $p<.0001$; /s/-/z/: $p<.0001$). With respect to individual differences, Mark shows an individual trend with a considerable overlap in the values for the /t/ and /d/ contrast. This is most likely due to Mark's position as a /t/ and /d/ flapper, which makes him partially neutralise the /t/-/d/ contrast, as described in the literature on North American English (Braver, 2011; Derrick & Gick, 2011; Zue & Laferriere, 1979 for American English). Finally, as shown in Figure 7, all five department members use voicing as a correlate of the fortis-lenis contrast.

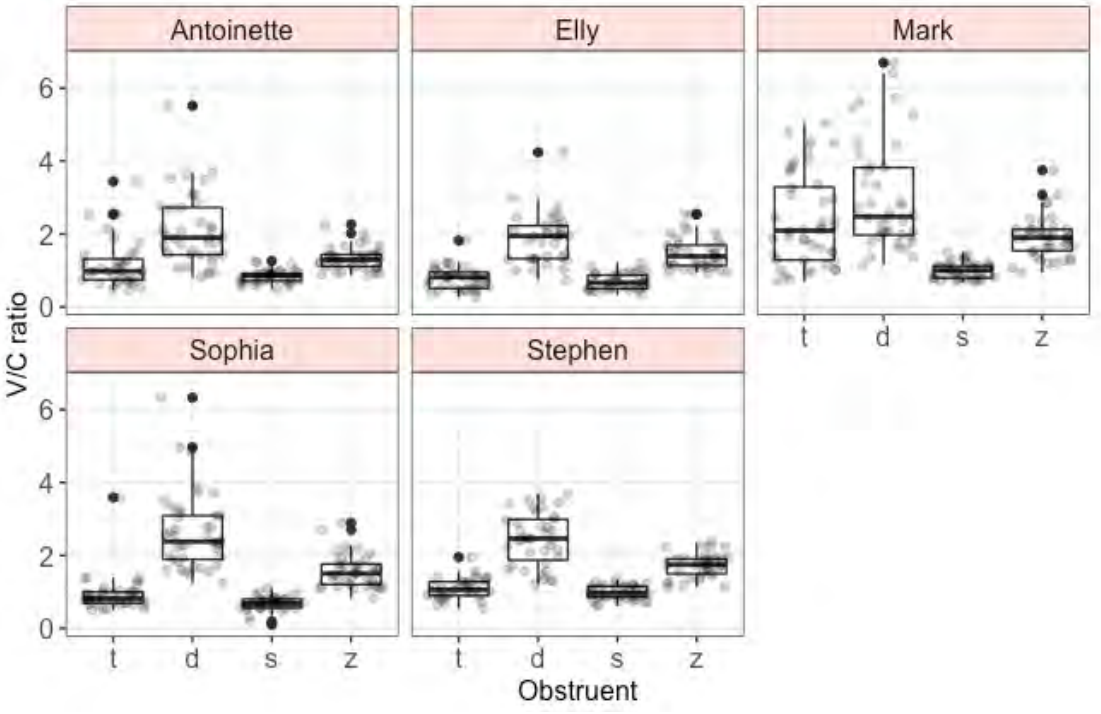


Figure 6. V/C ratio and the fortis-lenis contrast by individual.

Overall, the statistical analyses confirm that consonant type is a significant predictor of the presence of voicing ($F(3, 23)=13.412$, *SS*=17304, $p<.0001$).[11] Post-hoc Tukey tests show that the fortis-lenis pairs are strongly significant against each other (/t/-/d/: $p<.0001$; /s/-/z/:

[10] Final model: lmer(V/C proportion ~ consonant + (1|speaker) + (1|word), data=data).

[11] Final model: glmer(voicing ~ consonant + (1|speaker) + (1|word), family="binomial', data=data).

p<.0001) while /t/-/s/ are weakly significant against each other (p=.09), and /d/-/z/ are not (p=.13). However, there is clear individual variation in the specific amounts of voicing produced to maintain this distinction:

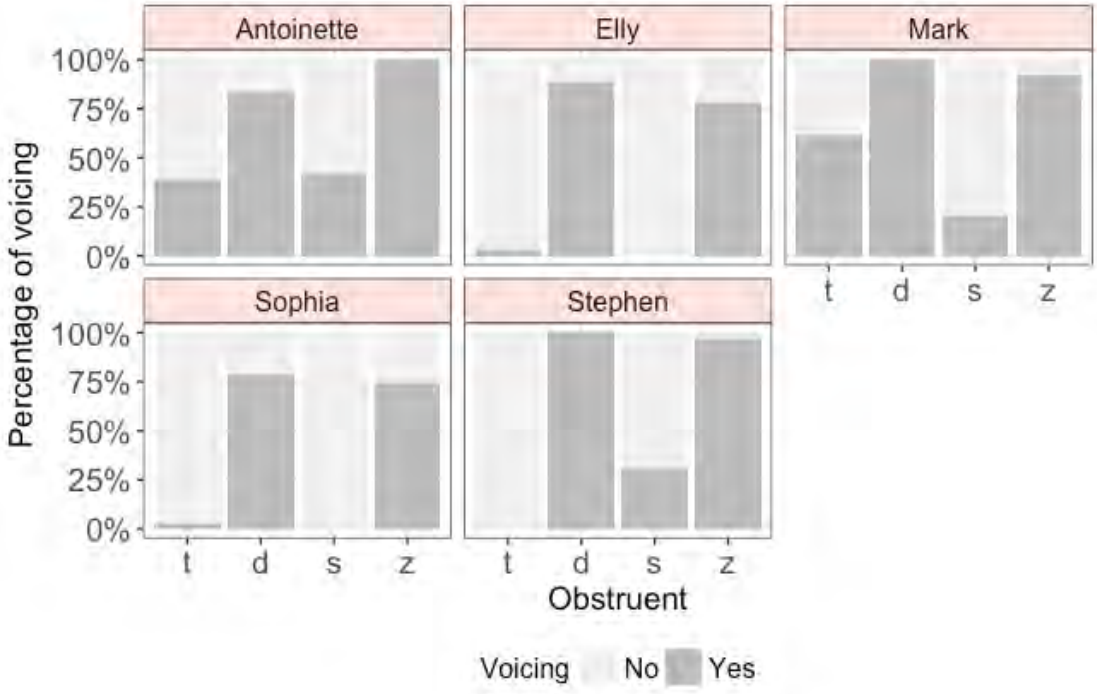


Figure 7. Presence of voicing (%) for each obstruent by individual.

Sophia and Elly are very consistent in that they never voice their /s/, whilst they produce voicing in 74% and 78% of their /z/ productions, respectively. Stephen is very consistent in that he never voices his /t/ and always voices his /d/. On the other hand, Stephen's fricatives show a less straightforward situation: whilst he voices both /s/ and /z/, he voices his /z/ more or less obligatorily (97%), whereas he voices his /s/ in 31%. Furthermore, the duration of the voicing in Stephen's /z/ reaches 47% of the overall consonantal duration on average, in contrast to that of 8% in his phonetically voiced /s/'s. Although Antoinette shows voicing in all of her obstruents, it is less frequent in her /t/ (39%) than /d/ (84%), and in her /s/ (42%) than /z/ (100%). Additionally, the duration of Antoinette's voicing is always shorter in the fortis obstruents (/t/: 18%; /d/: 39%; /s/: 7%; /z/: 21%). Mark presents a more complicated picture. His /d/ is categorically voiced and his /z/ is practically always voiced as well (93%); he voices his /s/ in 21% of the times and his /t/ in 61% of the times, which still renders him a speaker who utilises voicing to distinguish the fortis-lenis contrast in his production, albeit somewhat differently from the other department members.

3.2. Is the implementation of the contrast affected by prosodic variation?

This section highlights the influence of focus condition and contour type on the patterns shown in section 3.1. As shown in Figure 8, the duration of the pre-obstruent vowel in the implementation of the fortis-lenis contrast correlates with the type of intonation contour produced as follows: level contours (high, low) are associated with the shortest vowel durations; complex contours (rise-fall, fall-rise) are associated with the longest vowel durations and simple contours fall in between. This effect is highly significant ($F(5, 715)=29.83$, $SS=92191$, $p<.0001$). We found no other effects of contour type on the three correlates of the contrast.

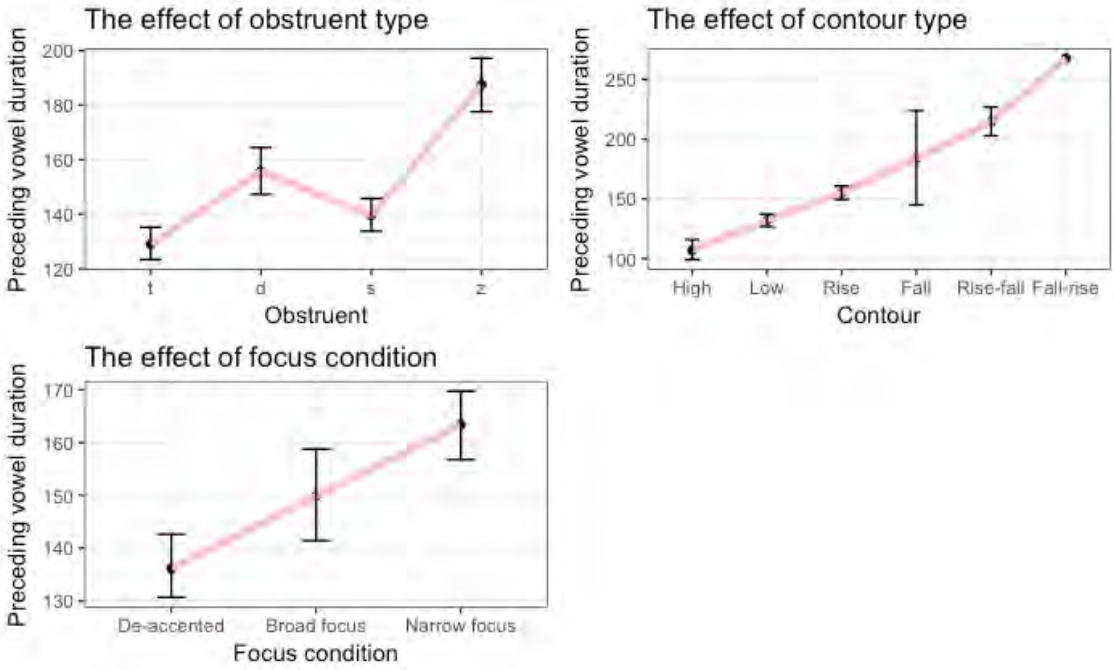


Figure 8. Effects of obstruent, contour and focus on pre-obstruent vowel duration (ms).

The role of the duration of the pre-obstruent vowel in the phonetic implementation of the fortis-lenis contrast is strongly affected by focus condition, with the shortest vowel durations being associated with the de-accented condition ($\bar{x}$=136.16ms, SD=44.62), the longest with the narrow focus condition ($\bar{x}$=163.42ms, SD=63.96), and broad focus in between ($\bar{x}$=149.89ms, SD=55.54). However, focus condition did not improve the model as a main effect. Its interaction with the type of consonant, on the other hand, was a significant predictor of the duration of the preceding vowel and was kept as part of the final model (F(6,

715)=12.875, *SS*=10662, *p*=.0089)[12]. Figure 9 illustrates this interaction. It can be seen from this graph that all four obstruents adhere to the trend for longer durations of the preceding vowel in the narrow focus condition described above, but that the main drivers of the trend are the voiced pair of obstruents, /d/ and /z/. Post-hoc Tukey tests reveal that the difference in length of the preceding vowel between de-accented and narrow focus is highly significant with /d/ (*p*=.0009) and /z/ (*p*=.008), while it is not significant with /t/ (*p*=.93) or /s/ (*p*=.98).

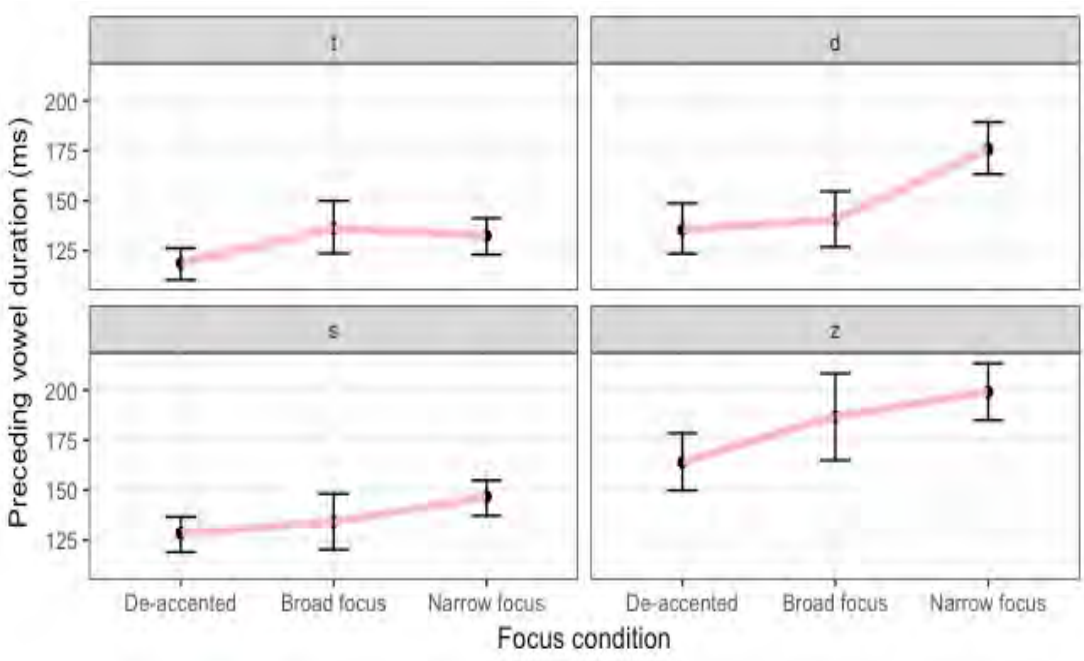


Figure 9. The interaction between obstruent and focus on pre-obstruent vowel duration (ms).

The regression model investigating V/C ratio and focus contains a significant interaction with consonant type (F(6,724)=3.6932, SS=8.4122, *p*=.0013)[13]. Tukey tests of the interaction show a magnification of the effect in the narrow focus condition (/t/-/d/: *p*<.0001; /s/-/z/: *p*<.0001) compared to the de-accented condition (/t/-/d/: *p*=.0002; /s/-/z/: *p*=.0015). These results are summarised in Figure 10.

[12] Final model: lmer(vowel duration ~ consonant * focus condition + intonational contour + (1|speaker) + (1|word), data=data).

[13] Final model: lmer(V/C ratio ~ consonant + focus condition + intonation contour + consonant * focus condition + (1 | speaker) + (1 | word), data = data).

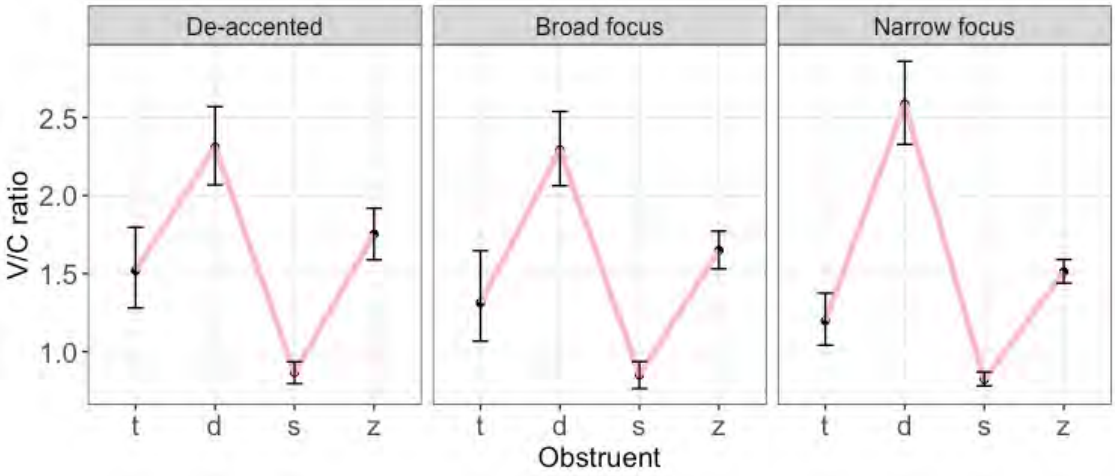


Figure 10. Focus effects on the preceding vowel duration (ms)

Regarding voicing, focus condition has a small, but non-significant, effect on its presence (Figure 11):

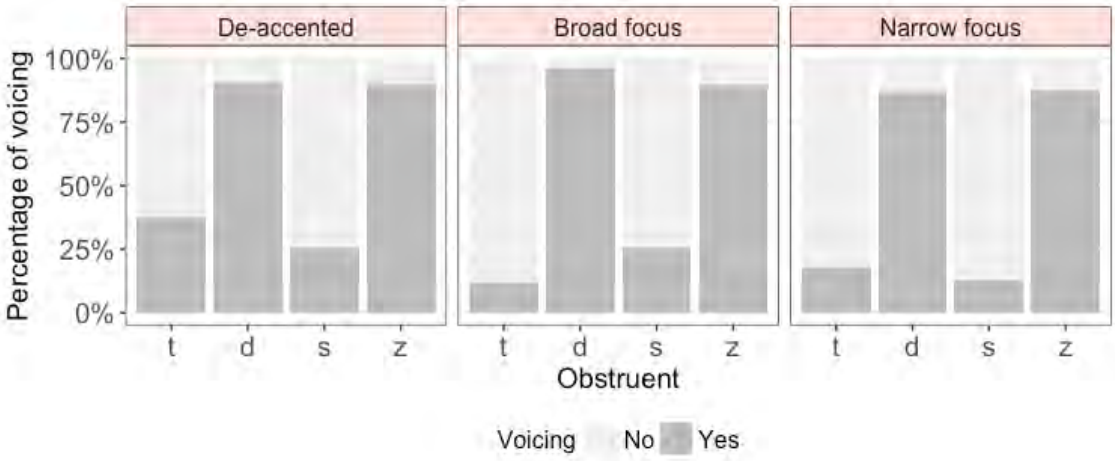


Figure 11. Presence of voicing (%) and the obstruent by focus condition.

On the whole, the fortis-lenis contrast distinction by voicing is more obvious in the broad focus condition as opposed to the de-accented condition, but this trend only approaches significance (*SS*=4148, *p*=.078) and does not improve the model. Interestingly, there is a significant interaction between obstruent and focus condition: the /t/-/d/ contrast is distinguished by the presence of voicing significantly more so in broad

and narrow focus than in the de-accented condition ($F(6,709)=2.3937$, $SS=6048.3$, $p=.0269$).[14]

In addition to the analyses presented so far, we also carried out a random forest analysis. This analysis involved the presence of voicing, the duration of the preceding vowel, and the V/C ratio as predictors. It revealed that, when it comes to determining the identity of the obstruent in question, V/C ratio is more important than the presence of voicing, which is in turn more important than the duration of the preceding vowel. Considering the /s-/z/ contrast, the factor most important for the identity of the obstruent is V/C ratio, followed by the presence of voicing, followed by vowel duration (voicing: 51.037; vowel duration: 24.414; V/C ratio: 102.141).[15] Regarding the /t/-/d/ contrast, the same tendency is observed: the most important factor determining the identity of the obstruent is V/C ratio, followed by the presence of voicing, followed by vowel duration (voicing: 59.923; vowel duration: 33.701; V/C ratio: 72.435). These results do not differ depending on whether the duration of the preceding vowel is normalised or raw.

3.3. Other correlates of the contrast

In this section, we offer a qualitative discussion based on descriptive statistics for potential correlates of the fortis-lenis contrast other than the preceding vowel duration, presence of voicing, and vowel/consonant durational ratio. Apart from the usually discussed correlates of the fortis-lenis contrast in English (as described in 3.1.-2.), we also report a range of other correlates of the contrast displayed by the individual department members.

Firstly, pre-aspiration is found consistently with all speakers only in their fortis obstruents[16], and always more frequently in /s/ than /t/ (Figure 12). Stephen and Mark are somewhat shy pre-aspirators.

---

[14] Final model: lmer(presence of voicing ~ consonant + focus condition + intonation contour + consonant * focus condition + consonant*intonation contour + (1 | speaker) + (1 | word), data = data).

[15] These numbers reflect the Gini importance index, which is used to refer to node impurity in the models (Louppe, Wehenkel, Sutera, & Geurts, 2013). The higher the number, the higher the misclassification of the obstruent is (as fortis vs lenis) if the variable of interest is left out. The more important variables will therefore show higher numbers.

[16] Sophia is a bit of an outlier in that she pre-aspirates her /d/ once and also her /z/ once.

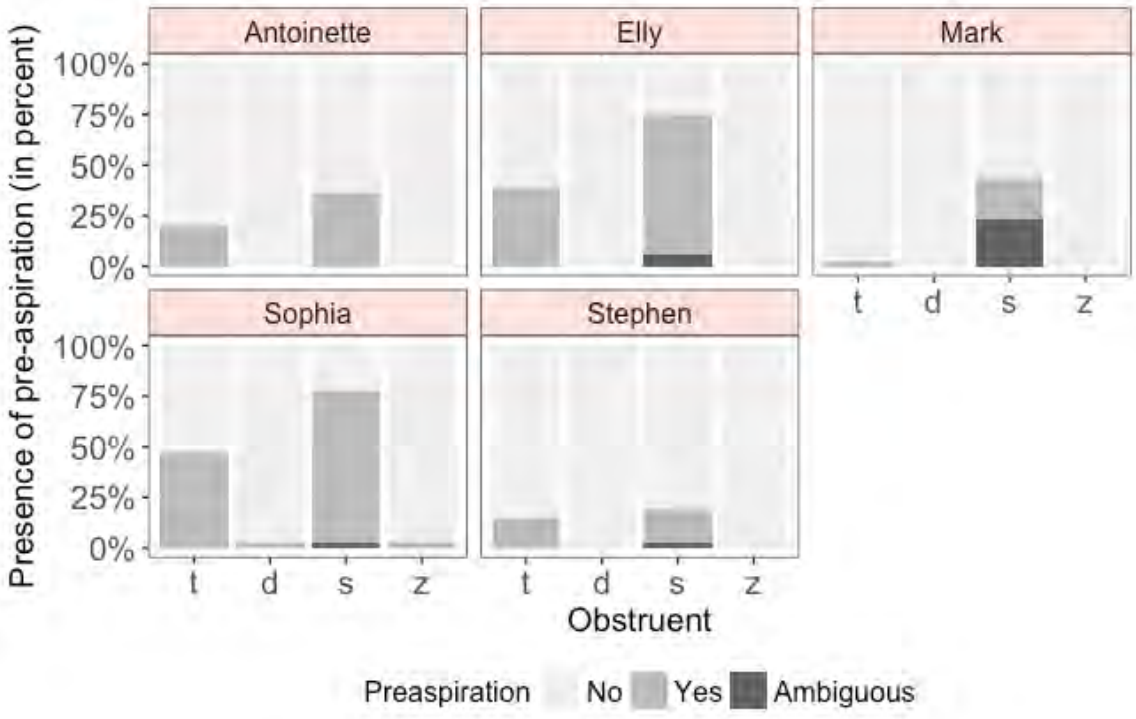


Figure 12. Presence of pre-aspiration (%) and the obstruent by individual.

As shown in Figure 13, affrication is clearly used as a correlate of the fortis-lenis /t/-/d/ contrast by Stephen, who never affricates his /d/'s, but affricates 94% of his /t/'s. Mark follows suit in that he never affricates his /d/'s, but his /t/'s are not as frequently affricated as Stephen's, reaching only 34%. Whist Sophia still gets to join the club of fortis plosive affricators in that she affricates her /t/'s more frequently than her /d/'s, the difference is far from clearcut in her case (/t/: 63%; /d/: 50%). Antoinette, on the other hand, shows a pattern unlike that of the three predominant /t/-affricators. Antoinette does affricate, but her affrication is higher in the context of /d/ than that of /t/ (/t/: 22%; /d/: 59%), which makes her a predominant /d/-affricator. As discussed further below, this is because her /t/'s are frequently spirantised. Finally, Elly also shows a unique behaviour by simply affricating both /t/ and /d/ more or less obligatorily, thus not utilising the feature as a correlate of the contrast at all.

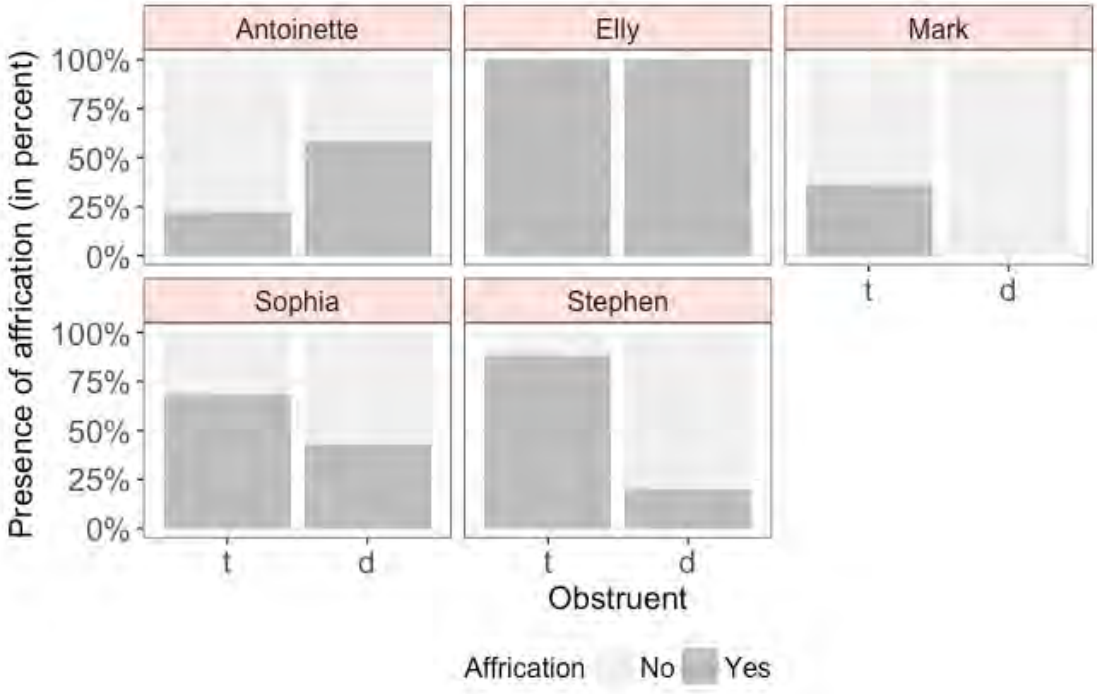


Figure 13. Presence of affrication (%) and the obstruent by individual.

If we consider the traditional descriptions of English as a post-aspirating language, looking into the use of post-aspiration as a potential correlate of the /t/-/d/ contrast in our five speakers presents us with a surprise: post-aspiration is fairly marginal in the dataset. Moreover, it does not consistently occur only in /t/, as illustrated in Figure 14, but also in /d/. Although Antoinette is the only "well-behaved" department member in that she post-aspirates only her /t/, it needs to be noted that even Antoinette does not quite present us with the canonical post-aspirated /t/ because, in her case, the post-aspiration is often found following a fully spirantised /t/ (see further below). To make the situation even more variable, Stephen presents us with some plosive releases which are ambiguous as to their being post-aspirated or unaspirated. Stephen and Mark are the only flappers in our dataset (Figure 15). Stephen is only a sporadic /d/-flapper (8%), while Mark flaps both his /t/'s (25%) and his /d/'s (28%), at a rate of application fairly comparably across his /t/ and /d/ categories. In contrast, the female members of the department do not flap. On the whole then, we can say that flapping is not a strategy used to distinguish the /t/-/d/ contrast by the five department members.

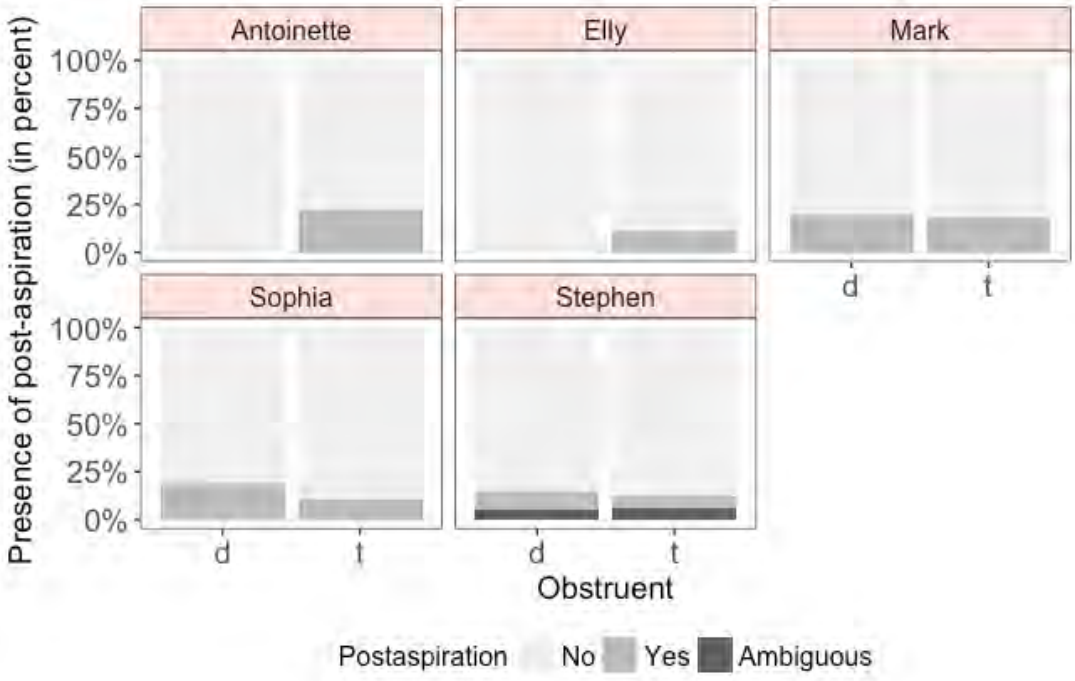


Figure 14. Presence of post-aspiration (%) and the obstruent by individual.

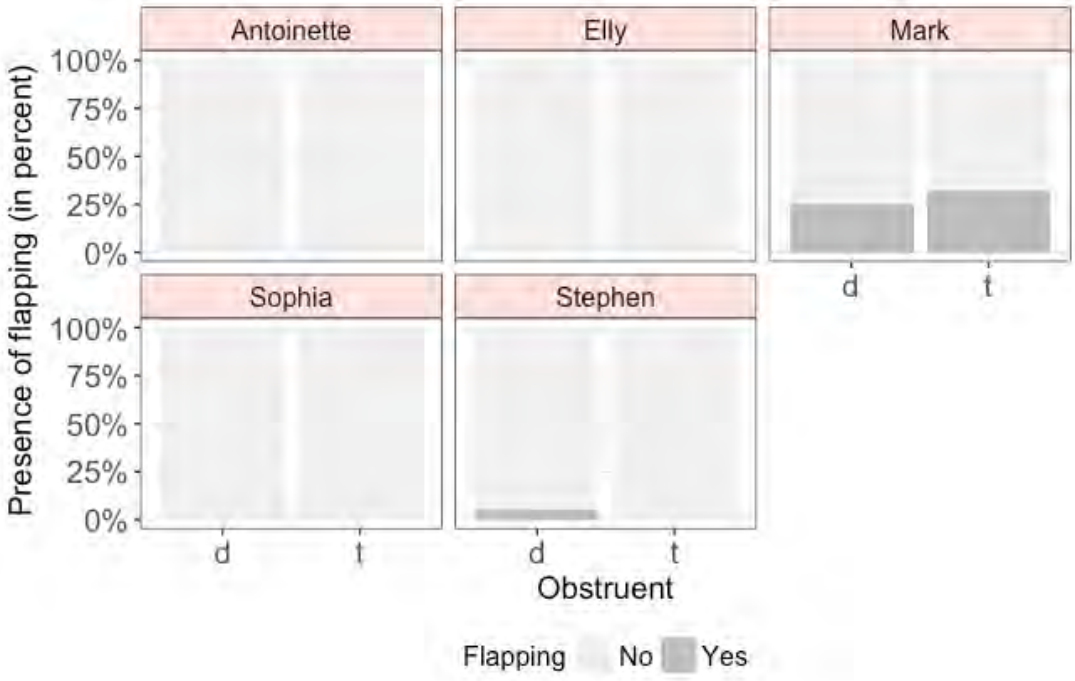


Figure 15. Presence of flapping (%) and the obstruent by individual.

As shown in Figure 16, Stephen is in the lead of the team of non-spirantisers, contrary to what has been claimed about /t/ spirantisation in southern Irish English dialects (Hickey, 2004). The rest of his closure-preserving team, Elly and Mark, do spirantise but only marginally so. Antoinette, on the other hand, very frequently spirantises both /t/ (78%) and /d/ (41%). Sophia also enjoys spirantisation in both obstruents (/t/: 37%; /d/: 10%). Importantly, both Antoinette and Sophia spirantise their /t/'s more frequently than their /d/'s. In addition, for both Antoinette and Sophia, full spirantisation is more frequently associated with their /t/'s and semi-spirantisation with their /d/'s.

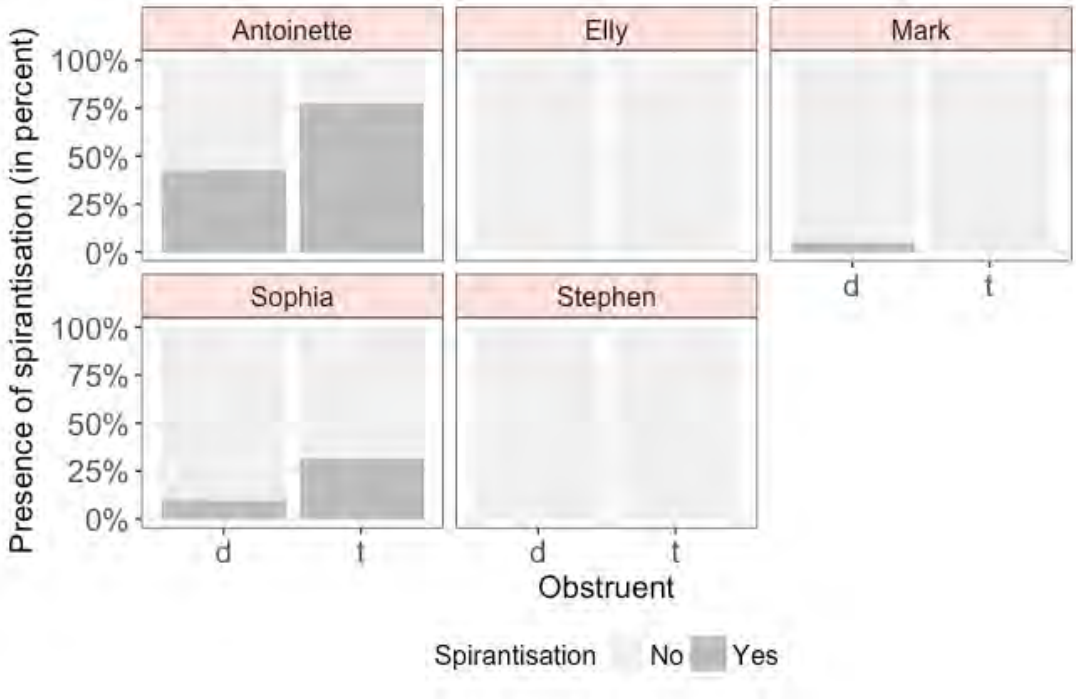


Figure 16. Presence of spirantisation (%) and the obstruent by individual.

3.4. Other correlates of the contrast and prosodic effects

As shown above, pre-aspiration contributes to the phonetic implementation of the fortis-lenis contrast in all five department members, albeit marginally so in some cases. Furthermore, Sophia, Stephen, Mark, and Antoinette use affrication as one of the correlates of the /t/-/d/ contrast. Finally, Sophia and Antoinette also employ spirantisation to distinguish /t/ and /d/. In this section, we look into whether these three features in the relevant speakers are subject to effects of focus and, if so, to what extent.

Firstly, as Figure 17 shows, pre-aspiration is the most frequent in the narrow focus condition, and this applies to all four obstruents. More specifically, /s/ is pre-aspirated 50% of the times in the focus condition, but only 34% of the times in both the broad focus and the de-accented conditions. In the same vein, /t/ is pre-aspirated 28% of the times in the narrow focus condition, which is higher than in the broad focus (21%) and the de-accented (18%) conditions.

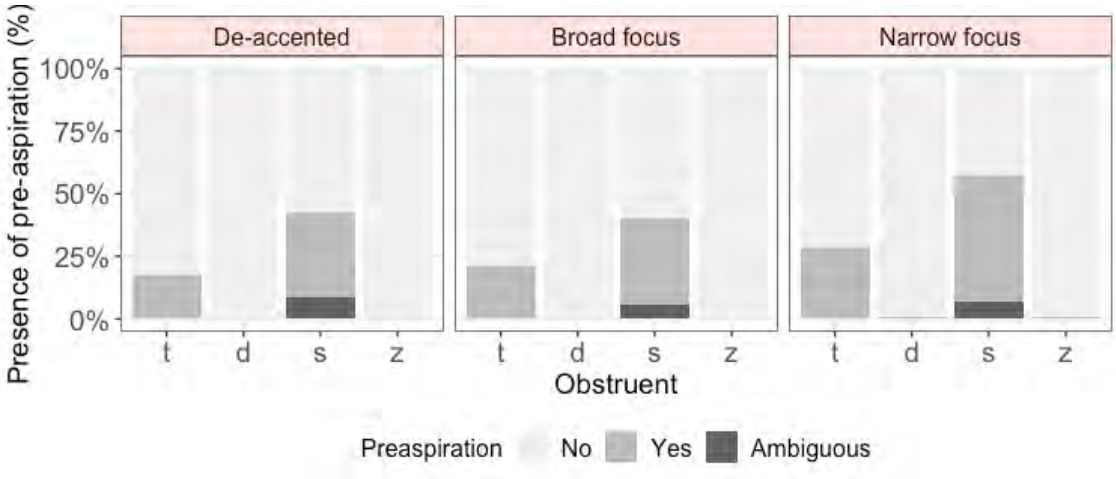


Figure 17. Presence of pre-aspiration (%) and the obstruent by focus condition.

Interestingly, the two cases of pre-aspirated lenis obstruents[17] are both found in the focus condition. On the whole, pre-aspiration application is a somewhat better correlate of the fortis-lenis contrast in the focus condition then; on the other hand, its marginal occurrence in the lenis series in this prosodic condition undermines this finding. The presence of affrication and that of spirantisation are not affected by focus condition in the relevant speakers.

4. Discussion and conclusions

This study has examined the effects of different levels of prosodic focus on the realisation of the fortis-lenis contrast, and on obstruent realisation more generally. Furthermore, we have taken a brief look at the contribution of f0 variation on the three main correlates of the fortis-lenis contrast under investigation: the duration of the preceding vowel, the voicing frequency, and the vowel/consonant ratio.

[17] As we have seen earlier, Sophia is the lenis obstruent pre-aspiration culprit.

Investigating these correlates on their own, we have found that V/C ratio is the most important of the three correlates in our data, followed by the presence of voicing. Our analyses of the duration of the pre-obstruent vowel did not yield significant results, although post-hoc tests revealed a significant pairwise difference between the duration of the vowel preceding /t/ as opposed to that preceding /z/, by which the duration was shortest before /t/ and longest before /z/. With respect to V/C ratio, there was an effect of consonant type whereby lenis obstruents had higher V/C ratios, i.e. the preceding vowel was proportionally longer than the consonant more so prior to lenis than fortis obstruents. In addition, the presence of voicing in the obstruent also serves as a correlate of the contrast.

When we added focus and the type of intonation contour to the models we saw that these have predictive effects on several of the correlates of the implementation of the fortis-lenis contrast. In formulating our first research question, we hypothesised that increased levels of focal prominence would correlate with higher reliability and availability of the consonantal correlates (MacWhinney, 2001, 2012), with a decrease as the focal prominence decreases. After considering the effect of focus on pre-obstruent vowel duration, we found a trend for shorter vowel durations in the de-accented condition, through mid-level durations in broad focus and with the longest durations in the narrow focus condition. This trend was not significant as a main effect; however, an interaction between focus and the type of consonant *was* significant. Closer scrutiny showed that the lenis set of obstruents was the main driver of the correlation between focus condition and vowel length. A similar significant interaction between focus and consonant type was reported for the V/C ratio correlate, and a non-significant interaction was also reported for the third correlate, namely the presence of voicing. In both these cases, the fortis/lenis distinction was magnified in the narrow focus condition as compared to the two other conditions. Our prediction is thus partly confirmed: different levels of focus as elicited in this study correlate at least partially with stronger manifestations of the fortis-lenis contrast with all three main correlates measured, although these correlations were not always statistically significant.

Considering other potential correlates of the contrast (pre-aspiration, post-aspiration, affrication, spirantisation, flapping, glot-talling, ejectivisation), we found that, firstly, all of the speakers use pre-aspiration to contribute to the fortis-lenis distinction (although a bit more strongly for the fricative rather than the plosive pair). Secondly,

affrication seems to contribute to the /t/-/d/ distinction for four of the speakers, with /t/ being affricated more frequently for three of these. Aspiration is only marginally employed to distinguish /t/ from /d/, and only in one speaker. Finally, spirantisation is higher for /t/ than /d/ in the two speakers who spirantise.

With regard to obstruent realisation, in connection with our second research question we predicted that increased levels of prosodic prominence would correlate with increasing amounts of variation in obstruent realisation, especially in the case of pre-aspiration. In our data, only pre-aspiration is sensitive to the different focus conditions in that it applies more frequently in the narrow focus condition than in the broad focus and the de-accented positions. This indicates that it is a somewhat more robust correlate of the contrast in the narrow focus condition. The fact that we find two cases of pre-aspirated lenis obstruents in the narrow focus condition may suggest that pre-aspiration as such is more likely to innovate in a more prosodically prominent condition. This fits in well with the overall findings of pre-aspiration being a laryngeal aspect of stressed syllables (Hejná, 2015, for an overview), and constitutes a partial confirmation of our hypotheses.

Finally, we briefly investigated the effects of different nuclear contours on obstruent realisation as formulated in our third research question. Our hypotheses include a tendency for increased contrast strength with high or rising intonational movements, as well as a correlation between shorter duration and level or simple contours (and between longer durations and complex contours). This has often been referred to in the literature (since English is a compression rather than a truncation language – Grabe, 1998; e.g. Ohala, 1978), but has not to our knowledge been attested through experimental work. We did not find any significant correlates between the type of nuclear contour and the presence of voicing or V/C ratio, but we did find a correlation with the duration of the preceding vowel whereby level contours (high, low) were associated with shorter vowel durations than simple contours (falls, rises), which were again associated with shorter vowel durations than complex contours (rise-falls, fall-rises). Our hypotheses stemming from the final research question are thus also partially fulfilled.

On the whole, then, there is a tendency for the correlates of the fortis-lenis obstruent contrast to be noticeably affected by prosodic focus and the complexity and directions of f0 movements. This suggests that prosody-segment interactions are a worthwhile avenue to pursue.

5. Implications for studies of focus, pitch contours and the fortis-lenis contrast

But how *exactly* should this avenue be explored? Firstly, the methodology used deserves further elaboration. As briefly reported in the results section, the finding that increased levels of focus affected the implementation of the fortis-lenis contrast to some extent, but did not yield many clear effects, may stem from difficulties in getting our department members to produce the intended levels of focus, from differences in the manifestations of narrow focus (and de-accentuation) produced by individual speakers, or from exaggerated or performed speech styles. While previous work has not tended to report on such difficulties, in our study different degrees of focus were generally realised in gradual rather than overtly categorical terms, and were highly speaker-specific. While we feel this probably reflects patterns found in spontaneous speech to some degree (something that needs to be confirmed in further research), our findings on focus need to be replicated and extended by further work.

Secondly, the general lack of effects of intonation contours on the realisation of consonants could in part be due to our lack of detailed acoustic measurements. The tendency of rising f0 to correlate with higher centres of gravity across fricatives (Niebuhr, 2012) indicates that such f0 movements may have effects on other aspects of frication, such as the intensity of the aspiration and the affrication noise components in obstruents and potentially other acoustic characteristics of their realisation as well. Another explanation might be that English is not a truncating language, unlike German, a fact which may limit the need to express the information carried by truncated intonational contours through consonantal means (as is the case in German). Further work could address these potential gaps by taking additional acoustic measures, such as spectral moments.

In addition, whilst the finding that vowel duration and the complexity of the intonational contour are correlated is in line with common wisdom in phonetics and phonology (see e.g. Ohala, 1978), to the best of our knowledge this has not in fact been demonstrated by experimental work for non-tone languages (and even with tone languages the picture is unclear, cf. Köhnlein, 2015, p. 232). Further work may therefore want to provide additional evidence for this finding using more controlled speech samples. In addition, high degrees of between-speaker differences suggest potential dialectal variation which could be fruitful

for future sociolinguistic work: there is more to consonantal variation than first meets the eye, and we conclude that it is not necessarily the case that vowels do in fact exhibit more variation than consonants.[18]

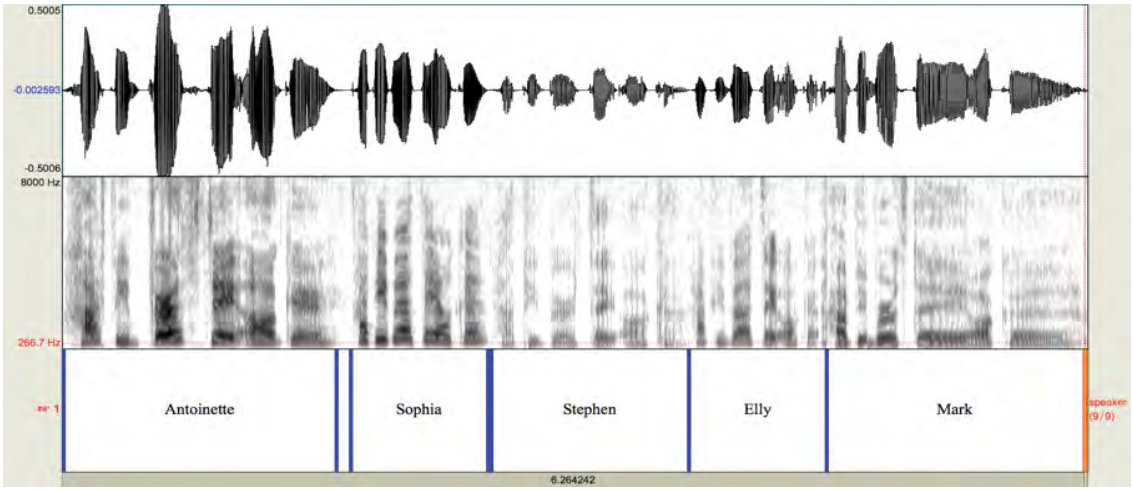
Finally, we note that it is not post-aspiration that distinguishes /t/ and /d/ foot-medially and -finally in our sample. For some speakers, affrication is employed as a correlate of the contrast, but this correlate is only a marginal strategy. Although we did not measure VOT, our observation of the data from the annotation of the other potential correlates suggests that VOT is an important correlate. If this is indeed the case, it is the release duration of the plosive (traditionally quantified as VOT) which is more important than whether the plosive is post-aspirated, affricated, both, or neither. This is in line with the findings available for Aberystwyth English (Hejná, 2016b).

Acknowledgements

We would like to thank our five departmental guinea pigs as well as Oliver Niebuhr for his very useful comments on an earlier draft of this paper.

Comments

Dearest Ocke, let your life be forever full of vowels, consonants, and prosodic phenomena. “Happy birthday, Ocke!” from us, but also from our five lovely participants: Antoinette, Sophia, Stephen, Elly, and Mark!



[18] However, it would be rather difficult to actually test this claim.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

Baumann, S., Becker, J., Grice, M. & Mücke, D. (2007). Tonal and articulatory marking of focus in German. *Proceedings of ICPhS 2007,* Saarbrücken*,* 1029-1032.

Boersma, P., & Weenink, D. (1992-2017). *Praat: doing phonetics by computer.* Version 6.0.25.

Baltazani, M., & Jun, S.-A. (1999). Focus and topic intonation in Greek. *Proceedings of ICPhS 1999*, San Francisco, 1305-1308.

Bohn, O.-S., & Caudery, T. (2017). *The sounds of English. An activity-based course in English phonetics and phonology.* Aarhus University Press, Aarhus.

Braver. A. (2011). Incomplete neutralization in American English flapping: production study. *University of Pennsylvania Working Papers in Linguistics* / *Proceedings of the 34th Annual Penn Linguistics Colloquium*, *17*(1), 29-40.

Cox, F., & Palethorpe, S. (2007). Australian English. *Journal of the International Phonetic Association*, *37*(3), 341-350.

Cruttenden, A. (1996) *Intonation*. Cambridge: Cambridge University Press.

Derrick, D., & Gick, B. (2011). Individual variation in English flaps and taps: a case of categorical phonetics. *The Canadian Journal of Linguistics* / *La Revue Canadienne de Linguistique*, *56*(3), 307-319.

D'Imperio, M. (2001). Focus and tonal structure in Neapolitan Italian. *Speech Communication*, *33*, 339-356.

Docherty, G. (1992). *The timing of voicing in British English obstruents*. New York: Foris Publications.

Dohen, M., Loevenbruck, H., & Hill, H. (2006). Visual Correlates of Prosodic Contrastive Focus in French: Description and Inter-Speaker Variabilities, *Proceedings of Speech Prosody 2006,* Dresden, 221-224.

Fox, J. (2003). Effect displays in R for Generalised Linear Models. *Journal of Statistical Software*, *8*(15), 1-27.

Fox, J., & Palethorpe, S. (2007). Illustrations of the IPA: Australian English. *Journal of the International Phonetic Association*, *37*(3), 341-350.

Gordeeva, O., & Scobbie, J. (2013). A phonetically versatile contrast: pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *Journal of the International Phonetic Association*, *43*, 249-271.

Grabe, E.. (1998). *Comparative Intonational Phonology: English and German*. MPI Series in Psycholinguistics 7, Wageningen, Ponsen en Looien.

Görs, K., & O. Niebuhr. (2012). Hocus Focus – How prosodic profiles of contrastive focus emerge and change in different elicitation contexts. *Proceedings of Speech Prosody*, Shanghai, China, 262-265.

Hejná, M. (2015). *Pre-aspiration in Welsh English: a case study of Aberystwyth* (PhD thesis), University of Manchester, Manchester.

Hejná, M. (2016a). Pre-aspiration: manual on acoustic analyses 1.1. Manuscript released on LingBuzz.

Hejná, M. (2016b). Multiplicity of the acoustic correlates of the fortis-lenis contrast: plosives in Aberystwyth English. *Proceedings of Interspeech 2016*, San Francisco, 3147-3151.

Hejná, M., & Kimper, W. (Submitted). Pre-closure laryngeal properties as cues to the fortis-lenis plosive contrast in British English. *Yearbook of the Poznań Linguistic Meeting 2017*.

Hickey, R. (2004). The phonology of Irish English. In B. Kortmann & E. W. Schneider (Eds.), *Handbook of varieties of English* (Vol. 1, *Phonology*, pp. 68-97). Berlin: Mouton de Gruyter.

Hua, C., Li, B., & Wayland, R. (this volume). *Native and non-native English speakers' assessment of nuclear stress produced by Chinese Learners of English*.

Iverson, G. K., & Salmons, J. C. (2006). On the typology of final laryngeal neutralisation: evolutionary phonology and laryngeal realism. *Theoretical Linguistics*, *32*(2), 205-216.

Jansen, W. (2004). *Laryngeal contrast and phonetic voicing: a laboratory phonological approach to English, Hungarian, and Dutch*. PhD thesis, University of Groningen, Groningen.

de Jong, K. J. (1995). The supraglottal articulation of prominence in English: linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*, *97*(1), 491-504.

Kingston, J. (1986). Are f0 differences after stops accidental or deliberate?. *111th Meeting of the Acoustical Society of America*, S27.

Kingston, J., Diehl, R. L., Kirk, C. J., & Castleman, W. A. (2008). On the internal perceptual structure of distinctive features: the [voice] contrast. *Journal of Phonetics*, *36*, 28-54.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*(5), 1208-1220.

Kohler, K. (1984). Phonetic explanation in phonology. *Phonetica*, *41*, 150-171.

Kohler, K. (2011). Communicative functions integrate segments in prosodies and prosodies in segments. *Phonetica*, *68*, 26-56.

Köhnlein, B. Thee complex durational relationship of contour tones and level tones. *Diachronica*, *32*(2), 231-267.

Kügler, F. (2008). The role of duration as a phonetic correlate of focus. *Proceedings of Speech Prosody 2008*, Campinas, Brazil, 591-594.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1-26.

Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

Ladefoged, P. (1968). *A phonetic study of West African languages*. (2nd edition). Cambridge: Cambridge University Press.

Leemann, A., Kolly, M.-J., Li, Y., Chan, R., Kwek, G., & Jespersen, A. (2016). Towards a typology of prominence perception: the role of duration. *Proceedings of Speech Prosody 2016*, Boston, MA.

Louppe, G., Wehenkel, L., Sutera, A., & Geurts, P. (2013). Understanding variable importances in forests and randomized trees. *NIPS' 13 Proceedings of the 26th International Conference on Neural Information Processing Systems*, vol. 1, Lake Tahoe, Nevada.

MacWhinney, B. (2001). The Competition Model: the input, the context, and the brain. In P. Robinson (Ed.), *Cognition and Second Language Instruction*, (pp. 69-90). Cambridge: Cambridge University Press.

MacWhinney, B. (2013). The logic of the unified model. In S. M. Gass & A. Mackey (Eds.), *The Routledge Handbook of Second Language Acquisition* (pp. 211-227). London/New York: Routledge.

Mennen, I. (2006). Phonetic and phonological influences in non-native intonation: an overview for language teachers. *Working paper WP-9*. Edinburgh.

Mücke, D., & Grice M. (2014). The effect of focus marking on supralaryngeal articulation – Is it mediated by accentuation? *Journal of Phonetics*, *44*, 47-61.

Niebuhr, O. (2008). Coding of intonational meanings beyond F0: evidence from utterance-final /t/ aspiration in German. *Journal of the Acoustical Society of America*, *124*(2), 1252-63.

Niebuhr, O. (2012). At the Edge of Intonation: The interplay of utterance-final F0 movements and voiceless fricative sounds. *Phonetica*, *69*(1-2), 7-27.

Niebuhr, O. (2013). Coding of intonational meanings beyond F0: evidence from utterance-final /t/ aspiration in German. *Journal of the Acoustical Society of America*, *124*, 1252.

Niebuhr, O. (this volume). Pitch accents as prosodic constructions – Evidence from pitch-accent specific micro-rhythms in German.

Niebuhr, O. & Michaud, A. (2015) Speech data acquisition - The underestimated challenge. *Kieler Arbeiten in Linguistik und Phonetik (KALIPHO)*, *3*, 1-42.

Niebuhr, O., Lill, C., & Neuschulz, J. (2011). At the segment-prosody divide: the interplay of intonation, sibilant pitch and sibilant assimilation. *Proceedings of the 17th ICPhS*, Hong Kong, 1478-1481.

Niebuhr, O., & Pfitzinger, H. (2010) On pitch-accent identification – The role of syllable duration and intensity. *Proceedings of Speech Prosody 2010*, Chicago, US, 1-4.

Nolan, F. (2006). Intonation. In B. Aarts & A. McMahon (Eds.), *Handbook of English linguistics* (pp. 433-457). Oxford: Blackwell.

Norcliffe, E., & Jaeger, T. F. (2005). Accent-free prosodic phrases? Accents and phrasing in the post-nuclear domain. *Proceedings of Interspeech 2005*.

Ohala, J. J. (1978). Production of tone. In V. A. Fromkin (Ed.), *Tone: a linguistic survey*, (pp. 5-39). New York: Academic Press.

Penney, J., Cox, F., Miles, K., & Palethorpe, S. (2018). Glottalisation as a cue to coda consonant voicing in Australian English. *Journal of Phonetics*, *66*, 161-184.

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

RStudio Team (2015). *RStudio: integrated development for R. RStudio, Inc.,* Boston, MA (http://www.rstudio.com/).

Sarkar, D. (2008). Lattice: multivariate data visualization with R. Springer, New York.

Scobbie, J. (2005). Interspeaker variation among Shetland Islanders as the long term outcome of dialectally varied input: speech production evidence for fine-grained linguistic plasticity. *QMUC Speech Science Research Centre Working Paper WP* 2.

Smith, C. (1997). The devoicing of /z/ in American English: effects of local and prosodic context. *Journal of Phonetics*, *25*(4), 471-500.

Smith, J. L. (2002). *Phonological augmentation in prominent positions* (PhD thesis). University of Massachusetts Amherst, Massachusetts.

Steriade, D. (1998). Alternatives to syllable-based accounts of consonantal phonotactics. *Proceedings of LP '98*, 205-245.

Stevens, M., & Hajek, J. (2005). Spirantization of /p t k/ in Sienese Italian and so-called semi-fricatives. *Proceedings of Interspeech 2005*, Lisbon, 2893-2897.

Su, V. W. Y. (2007). *The gender variable in Australian English stop consonant production* (BA thesis). The University of Melbourne, Melbourne.

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.

Trousdale, G. (2010). *An introduction to English sociolinguistics*. Edinburgh University Press, Edinburgh.

Turk, A. E., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, *35*, 445-457.

Wagner, P., Origlia, A., Avesani, C., Christodoulides, G., Cutugno, F., D'Imperio, M., Escudero Mancebo, D., Gili Fivela, B., Lacheret, A., Ludusan, B., Moniz, H., Ní Chasaide, A., Niebuhr, O., Rousier-Vercruyssen, L., Simon, A.-C., Šimko, J., Tesser, F., & Vainio, M. (2015) Different parts of the same elephant: a roadmap to disentangle and connect different perspectives on prosodic prominence. *Proceedings of the 18th International Congress of Phonetic Sciences*. University of Glasgow, Glasgow.

Wells, J. C. (1990). Syllabification and allophony. In S. Ramsaran (Ed.), *Studies in the pronunciation of English, a commemorative volume in honour of A. C. Gimson*, (pp. 76-86), New York: Routledge.

Wright, M. N., & Ziegler, A. (2017). ranger: a fast implementation of Random Forests for high dimensional data in C++ and R. *Journal of Statistical Software*, *77*(1), 1-17.

Xu, Y., & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, *33*, 159-197.

Zue, V. W., & Laferriere, M. (1979). Acoustic study of medial /t, d/ in American English. *Journal of the Acoustical Society of America*, *66*, 1039-1050.