Routledge
Taylor & Francis Group

🔓 OPEN ACCESS

No place for /h/: an ERP investigation of English fricative place features

Kevin Schluter[a] 🔟, Stephen Politzer-Ahles[b,c] 🔟 and Diogo Almeida[a] 🔟

[a]Psychology Program, Division of Science, New York University Abu Dhabi, Abu Dhabi, UAE; [b]Faculty of Linguistics, Philology & Phonetics, University of Oxford, Oxford, UK; [c]NYUAD Institute, New York University Abu Dhabi, Abu Dhabi, UAE

ABSTRACT
The representational format of speech units in long-term memory is a topic of debate. We present novel event-related brain potential evidence from the Mismatch Negativity (MMN) paradigm that is compatible with abstract, non-redundant feature-based models like the Featurally Underspecified Lexicon (FUL). First, we show that the fricatives /s/ and /f/ display an asymmetric pattern of MMN responses, which is predicted if /f/ has a fully specified place of articulation ([Labial]) but /s/ does not ([Coronal], which is lexically underspecified). Second, we show that when /s/ and /h/ are contrasted, no such asymmetric MMN pattern occurs. The lack of asymmetry suggests both that (i) oral and laryngeal articulators are represented distinctly and that (ii) /h/ has no oral place of articulation in long-term memory. The lack of asymmetry between /s/ and /h/ is also in-line with traditional feature-geometric models of lexical representations.

Introduction

Any theory of language processing needs to specify its theoretical commitments about the nature of the long-term memory representations that are used in language comprehension and production. In the case of speech perception, there is little controversy that the speech signal is influenced by idiosyncratic, speaker-dependent information, and that few reliable invariant acoustic cues (Perkell & Klatt, 2014) exist to allow a direct mapping between the signal and the seemingly abstract sound categories that human languages use for lexical storage. For instance, the acoustic information that is eventually associated with the consonant /t/[1] varies as a function of the vowel that follows it, and no two utterances of the same word are acoustically the same, even when pronounced by the same speaker (e.g. Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Newman, Clouse, & Burnham, 2001). Furthermore, the consonant /t/ is articulated differently in words like *stop* (unaspirated [t]), *top* (aspirated [tʰ]), *cat* (unreleased [t̚]), *water* (flap [ɾ]), and *button* (debuccalized [ʔ]), yet English speakers ultimately map them all to a more abstract category, which is /t/ in *phoneme*-based models of speech perception. Phonemes are specified at a more abstract level than simple acoustic or articulatory codes (see Baković (2014) for a recent discussion of what phonemes are and are not in speech perception).

Behavioural and theoretical research has shown that phonemes, and their constituent elements, are a useful construct in understanding speech processing and its relationship with lexical storage as well as acoustic and articulary correlates (Baković, 2014). Crucially, phonemes are generally characterised as bundles of sub-units called *features* (Chomsky & Halle, 1968; Halle, 1959), which receive abstract definitions conventionally labelled in terms of articulation. In this paper, we explore the predictions of a feature-based theory of phonology and speech perception. Features minimally distinguish two phonemes: for example, the stop consonants /t/ and /p/ differ only in the location of the closure in the mouth (place of articulation) and they are more similar to each other than /t/ and /m/, which differ in both articulatory place and the presence of nasal airflow (i.e. manner of articulation). The extent to which abstract, feature-based representations are encoded in the brain is still an open question; the present paper explores the neural long-term (or lexical) representation and organisation of features in English fricatives.

Phonological features and lexical storage

In feature-based theories, most consonants can be distinguished by using a few kinds of features, including place of articulation, manner of articulation, nasality, and voicing (laryngeal gesture). In this study, we focus

CONTACT Kevin Schluter ✉ kevin.schluter@nyu.edu

largely on place of articulation. Beyond describing the way sounds are articulated, features also play a role in the mental organisation and behaviour of speech sounds: many sounds produced with similar articulations or features (natural classes) undergo similar alternations. For example, English voiceless stops – /p/, /t/, and /k/ – are aspirated (pronounced with a puff of air) syllable-initially, but unaspirated when preceded by /s/: *pit~spit*, *take~stake*, *kit~skit*.

Features, then, are assumed to be a primary element in speech perception (and production). Nevertheless, it is not necessarily the case that all articulatory features need to be stored in lexical representations. For instance, while [t] may consist of a [Coronal] place, [Plosive] manner, and a [Spread Glottis] laryngeal gesture at the phonetic level, it may be unnecessary to store all these features in long-term memory when they are redundant. Rather, in a language like English, whenever the articulatory system receives commands to produce the features [Plosive] and [Spread Glottis], the place of articulation feature [Coronal] could be a default place feature that is inferred when neither [Labial] or [Dorsal] features are specified. Coronal sounds, then, can be said to be *underspecified* (Archangeli, 1988) with respect to their place-of-articulation feature in the lexicon (i.e. long-term memory). While a segment /t/ might be underspecified at a phonological level and lack a place feature, its phonetic correspondants (e.g. aspirated [tʰ] in *top*, unaspirated [t] in *stop*, or the flap [ɾ] in *water*) are specified [Coronal] when they are produced and initially perceived. When the articulatory system gets the command to produce a /t/, the missing feature is filled in with [Coronal] by default. In this way, underspecified features can be realised as gestural commands during articulation.

This kind of system may be more computationally efficient (Archangeli, 1984, 1988) for the process of perception and production of speech. There is ample evidence from synchronic sound alternations, especially assimilation patterns, that features like [Coronal] may be defaults (Avery & Rice, 1989). For instance, coronal sounds often assimilate to the place of articulation of following consonants (*rain + bow → rai*[m]*bow*) but non-coronals do not assimilate to coronals (*big + day → bi*[g]*day*, not *bi[d]*day*) or to other non-coronals (*book + bag → boo*[k]*bag*, not *boo[p]*bag*).

One implementation of this two-level feature-based system, where speech sounds can be fully characterised articulatorily but stored in long-term memory with only a subset of these observed properties, is the Featurally Underspecified Lexicon (FUL; Lahiri & Reetz, 2002, 2010). According to FUL, the parser attempts to fit the features of incoming sounds with categorical phonological representations in the lexicon. Because a phoneme can be underspecified, FUL utilises a ternary matching logic. An incoming sound may directly *match* an abstract representation when the features extracted from the speech signal map perfectly to the abstract representation. Sounds may also *mismatch* an abstract representation when the features extracted from the acoustic signal differ from those of an underlying representation. Finally, a *nomismatch* occurs when the features extracted from the input are consistent with an underlying phoneme but, because the underlying phoneme is not specified for the feature in question, they cannot be said to *match* per se. While there are some theoretical challenges to theories of underspecification (e.g. McCarthy & Taub, 1992; Mester & Itô, 1989), underspecification correlates with notions of markedness[2] and FUL provides clear and testable predictions.

Neurophysiological data from MMN studies and phonological features

In addition to synchronic sound alternations, FUL also receives support from neurophysiological data in the mismatch negativity (MMN) paradigm (Eulitz & Lahiri, 2004). The MMN is an early, fronto-central, negative-going event-related brain potential (ERP) component that emerges upon hearing a different stimulus embedded in a series of categorically identical stimuli (an oddball paradigm). The MMN peaks around 150–250 ms after the onset of a deviation from the repetitive signal. These repetitive, categorically identical stimuli, called *standards*, are thought to activate a stored representation, and the MMN is elicited when a *deviant* stimulus – incongruent with the representation of the standard – is encountered (for an overview, see Näätänen, Paavilainen, Rinne, & Alho, 2007). The amplitude of the MMN response correlates with the physical or perceptual distance between the standard and deviant stimuli, and the MMN has been shown to also reflect higher-level cognitive constructs like phonological category boundaries (Kazanina, Phillips, & Idsardi, 2006; Phillips et al., 2000).

The MMN for speech is sometimes asymmetrically sensitive to changes between the exact same sound categories (Eulitz & Lahiri, 2004) – when two sounds are presented in an oddball paradigm, the MMN may be larger when one is the deviant than when the other is. Crucially, these cases often involve underspecification.[3] FUL indeed predicts that the very same acoustic stimulus may be preferentially interpreted at two different levels of representation depending on whether it is presented as a standard or as a deviant in an MMN design.

This paradigm can thus reveal whether the lexical representation of a sound is different from its articulatorily

specified phonetic representations. For example, if the dorsal stop /k/ and coronal stop /t/ were both fully specified for their place of articulation ([Dorsal] and [Coronal], respectively) in long-term memory, then all else being equal, a deviant [t] in the context of a standard /k/ and a deviant [k] in the context of a standard /t/ would be equally different from one another, and therefore elicit MMNs of similar magnitude. If, however, [Coronal] is an underspecified feature in long-term memory, then FUL would predict that, in the context of a standard /t/, which primarily activates an *underspecified* long-term memory representation, a fully specified phonetic [k] is a *less different* stimulus (a *nomismatch*) than a fully phonetically specified deviant [t] is in the context of a fully specified standard /k/ (a *mismatch*). Thus, a deviant [k] in a standard /t/ context should elicit a smaller MMN than a deviant [t] in a /k/ context. This pattern has indeed been observed in a number of studies for coronal segments (Cornell, Lahiri, & Eulitz, 2013; Scharinger, Bendixen, Trujillo-Barreto, & Obleser, 2012) and voicing (Hestvik & Durvasula, 2016). While there may be other factors that can cause MMNs to be asymmetrical in these cases, underspecification theory makes a priori predictions about asymmetries across a variety of segments; in the Discussion we will examine other acoustic and phonetic accounts of asymmetries in light of our results.

These asymmetric MMN effects are challenging for other theories of lexical storage. Episodic theories grounded purely in acoustic memory (Goldinger, 1998) and feature-based theories that propose lexical storage with full articulatory specification (though see Mitterer, 2011) have no obvious mechanism with which to explain why only a specific subset of sounds seems to elicit asymmetric MMN effects. On the other hand, non-redundant, feature-based models like FUL not only propose the *nomismatch* as a natural explanation for these asymmetries, but also provide non-trivial and testable predictions.

Recent work on asymmetric MMN effects in the FUL framework has focused almost exclusively on place of articulation for vowels and stop consonants (Cornell, Lahiri, & Eulitz, 2011; Scharinger & Lahiri, 2010; Scharinger, Monahan, & Idsardi, 2012; Walter & Hacquard, 2004), where the predicted asymmetries for [Dorsal] and [Coronal] features have been replicated. However, the empirical findings for fricative contrasts are less clear. Bonte, Mitterer, Zellagui, Poelmans, and Blomert (2005) contrasted /f/ (phonetically and phonologically specified for [Labial] place of articulation) and /s/ (phonetically specified, but phonologically underspecified, for [Coronal]) embedded in nonwords. They found that the asymmetry predicted by FUL (larger MMN when [s] is the deviant) was only borne out when the fricatives

were embedded in contexts where they differed in phonotactic probability (contextual frequency), whereas an asymmetry in the opposite direction was found when phonotactic probability was controlled. They argue that the asymmetric MMN is due to phonotactic probability rather than feature underspecification (see also Tavabi, Elling, Dobel, Pantev, and Zwitserlood (2009) and Mitterer (2011) for similar arguments). Hestvik and Durvasula (2016) find an asymmetry for voicing, but the use of stop consonants forces the use of /ta/ and /da/ syllables, which may have introduced effects of phonotactic probability. Likewise, Gow (2001, 2002, 2003) and others have argued that underspecification is unnecessary for coronal place assimilation of stops and nasal consonants based on cross-modal priming in sentence contexts. The present study, however, focuses on fricatives presented in isolation, removing the possibility of photactic probability or phonological context for the segments. We examine the possibility of unigram frequency and find it cannot explain our result (see Table 1).

Furthermore this study uses segments which do not canonically alternate with each other in English: we contend the only mechanisms at play can be the acoustics of the segments themselves and their mental representations.

Goals and predictions of the study

The first aim of this study is to clarify whether fricatives show underspecification-related MMN asymmetries when the influence of phonological context or contextual frequency information is ruled out. We do this by presenting /f/ and /s/ in isolation – similar to vowels, which are also generally presented in isolation (Eulitz & Lahiri, 2004). FUL predicts that, without the influence of phonotactic frequency, the specified deviant in the context of the underspecified standard (deviant-[f] with standard-/s/) should yield a smaller MMN than the reverse case (deviant-[s] with standard-/f/). We test this in two experimental blocks (the FS blocks), one in which deviant-[s] with standard-/f/ is presented and the other in which deviant-[f] with standard-/s/ is presented (see Figure 1).

The second aim of the study is to use ERP data to refine the theory of phonological features. Unlike the well-studied oral segments, the featural makeup of glottal segments like /h/ has never been tested using this method. Moreover, while [Coronal] is considered to be a default feature, glottal segments are also sometimes considered to be a type of default, especially under feature geometry (e.g. Clements, 1985; Goldsmith, 1981; McCarthy, 1988). Though Lahiri and Reetz (2010)

Table 1. Segmental frequency in English. Log frequency values for both type and token frequency for /s/, /f/, and /h/.

Segment	Type frequency (log)	Token frequency (log)
/s/	23,068 (4.36)	175,792.23 (5.24)
/f/	5933 (3.77)	72,241.2 (4.86)
/h/	2781 (3.44)	87,684.83 (4.94)

suggest a [RADICAL] (i.e. tongue root) articulator feature for /h/,[4] /h/ is often represented with a laryngeal gesture feature like [SPREAD GLOTTIS], which may be represented on a distinct plane or tier from place features in the oral cavity (oral features) (cf. Avery & Idsardi, 2001). Under this representational format (also compatible with FUL's assumed features), laryngeal features should not conflict with a place feature such as [CORONAL].[5] The present study compares the glottal fricative /h/ to the coronal fricative /s/ in two experimental blocks (the HS blocks). In one of these blocks, deviant-[h] with standard-/s/ is presented, while deviant-[s] with standard-/h/ is presented in the other.

For the comparison between /h/ and /s/, four MMN patterns are logically possible, as shown in Figure 1. The first possible pattern illustrates the case in which /h/ is less specified in terms of its featural content than /s/: it may be the case that /s/ has some feature other

than [CORONAL] (possibly related to stridency or obstruency) that it shares with /f/, but that would clash asymmetrically with the features of /h/. In this case, deviant-[h] should yield a larger MMN than deviant-[s]. The second possibility is an MMN asymmetry similar to the one predicted by FUL for the FS comparison: A larger MMN when [s] is the deviant would be predicted if /h/ has a [RADICAL] articulator that is similar to the other oral articulators ([LABIAL] or [DORSAL]). In this case, the [CORONAL] feature of deviant-[s] should clash with the stored [RADICAL] feature of /h/, yielding a larger MMN. The third possible pattern is that both deviants elicit large and symmetrical MMNs (i.e. as large as that elicited by deviant-[s] with standard-/f/ in the FS blocks). Such a pattern would indicate that /h/ and /s/ have grossly incompatible featural specifications like /z/ and /n/ (Cornell et al., 2013) because a larger MMN is associated with a greater degree of difference (Sams, Paavilainen, Alho, & Näätänen, 1985). In other words, this pattern would suggest that /h/ is not a fricative in the same way as /s/, evidenced by the clash of features from /h/ to [s] and /s/ to [h]. The fourth possible pattern is that both deviants might elicit small and symmetrical MMNs (i.e. as small as that elicited by deviant-[f] with standard-/s/ in the FS blocks). This would be predicted if both

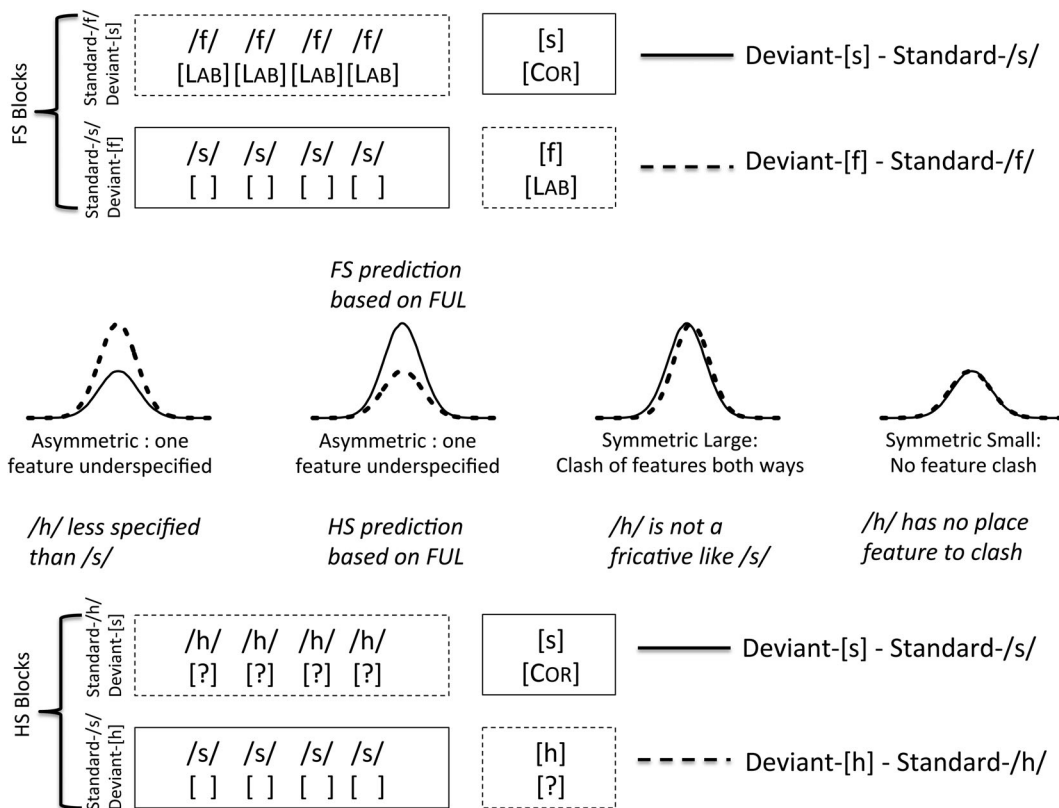
Figure 1. Conditions and predictions. The difference waves for the FS blocks (top) predict an asymmetric MMN pattern. The MMN pattern produced by difference waves for the HS blocks will be used to determine the featural representation of /h/.

segments *nomismatch* one another due to no clash of features. Formally, this would be the case if /h/ had no place of articulation feature at all, assuming that /s/ and /h/ are otherwise identical, a prediction that is compatible with some feature geometry proposals (Goldsmith, 1981). Given the current design, it may not be possible to distinguish the two symmetrical possibilities. Since the contrast between /f/ and /s/ is acoustically different from that between /h/ and /s/, they may not be directly comparable; therefore, if /h/ and /s/ do not elicit an asymmetrical MMN effect, there would be little basis for arguing that each of their MMNs is relativly large or that each is relatively small. Given this caveat, it is still fruitful to compare symmetry to asymmetry, and examine the direction of asymmetry.

Thus, the experiment follows a 2-by-2 factorial design manipulating CONTRAST (FS blocks contrasting /f/ and /s/, versus HS blocks contrasting /h/ and /s/) and DEVIANTSEGMENT (in two blocks, [s] was the deviant and the standard was either /f/ or /h/, depending on which contrast was being tested; in the other two blocks, /s/ was the standard and the deviant was either [f] or [h]). An asymmetry between /s/ and other phonemes (e.g. the expected asymmetry in FS blocks, shown in Figure 1) would be represented by an effect of DEVIANTSEGMENT in one or both contrasts. For each contrast, the MMN was calculated by measuring the ERP elicited by deviant tokens and subtracting from it the ERP elicited by the same tokens when used as a standard.

Materials and methods

Subjects

Twenty-six right-handed subjects (9 female, 18–23 years old, mean 20.5) from the NYU Abu Dhabi community participated in this study. Two were excluded because of excessive artefacts or technical problems, so data from a total of 24 subjects were analysed. All learned English at home before primary school (seven subjects reported being early bilinguals, acquiring French, Croatian, Norwegian, Hindi, Mandarin, Spanish, and Arabic concurrently alongside English), and reported normal hearing and no history of cognitive or linguistic impairment. All subjects provided informed consent and were paid for their participation. All methods for the study were approved by the Institutional Review Board of New York University Abu Dhabi.

Stimuli

The stimuli comprised five tokens each of [f], [s], and [h], recorded by one female native speaker of English in a sound-attenuated room (the stimuli are included in Supplementary File 1). Stimuli were recorded using an Electro-Voice RE20 cardioid microphone, and digitised at 22050Hz with a Marantz Portable Solid State Recorder (PMD 671). There were no surrounding vowels. Each fricative was shortened to 300 ms by cropping out a portion of the middle of the token. Tokens were not ramped, but natural onsets and offsets – cropped at zero-crossings – were used to avoid the perception of unintended stop consonants. Amplitude was normalised to 71 dB (see the Discussion for discussion of the potential consequences of this normalisation for MMN asymmetries). One reviewer pointed out that, phonetically, the [h] tokens used in the present study sound like voiceless vowels, in which case they may necessarily include coarticulation with a neutral vowel; we return to this point in the Discussion.

In order to assess the acoustic similarity of the experimental materials, each waveform was converted into matrices of 12 Mel-frequency cepstral coefficients (MFCCs) and 12 delta coefficients based on 15 ms windows spaced into 5 ms intervals (see supplementary materials for a larger and more in-depth acoustic analysis of the stimuli). MFCCs provide a useful and psychophysically motivated representation of the spectral characteristics of each stimulus at different points in time, and the delta coefficients characterise their spectral changes over time. Both are routinely used as input to measures of the similarity of speech sounds (e.g. Mielke, 2012). The resulting two matrices containing the spectral information for each stimulus (MFCCs and delta coefficients) were compared with the matrices resulting from every other stimulus, using the dynamic time warping algorithm (Mielke, 2012). These analyses generated two distance matrices, one based on MFCCs, representing time-frequency content, and one based on the delta coefficients, representing spectral change over time. The distance matrices were submitted for cluster analyses using the Neighbor-Joining Tree Estimation algorithm (Mielke, 2012; Saito & Nei, 1987). In addition, the two distance matrices were also subjected to DISTATIS, which is a generalisation of multidimensional scaling allowing the joint analysis of two or more distance matrices (Abdi, Williams, Valentin, & Bennani-Dosse, 2012). DISTATIS produces an optimal (in a least-squares sense) compromise space between two or more distance matrices, and the results can be visualised in a manner akin to the results of multidimensional scaling (see supplementary materials for more details and information). The output of the cluster and DISTATIS analyses is presented in Figure 2. As these analyses demonstrate, the stimuli can be categorised into their abstract phonemic categories on the basis of their acoustic properties

alone. Furthermore, the tokens from the /s/ and /h/ categories are the most acoustically dissimilar, with /f/ falling somewhere in the middle of the two categories; thus, an EEG result showing a larger MMN for the contrast between /f/ and /s/ (acoustically closer) than for the contrast between /h/ and /s/ (acoustically more distant) would not be predicted on the basis of acoustic properties alone. Likewise, frequency analysis of individual segments based on the IPhOD database (Vaden, Halpin, & Hickok, 2009) clarifies that /s/ is the most frequent segment of the three, and that /f/ and /h/ are similar to one another in terms of token and type frequency (see Table 1).

Procedure

The electroencephalogram was recorded during a passive oddball paradigm. The experiment consisted of four blocks. Two were FS blocks, comprising a deviant-[f] with standard-/s/ block and a deviant-[s] with standard-/f/ block. The other two were HS blocks: deviant-[h] with standard-/s/ and deviant-[s] with standard-/h/. Each block included 850 standard and 150 deviant trials. Stimuli were presented with a 700–1200 ms jittered interstimulus interval. Presentation order was pseudorandomized such that 2–10 standard stimuli intervened between each deviant, and each block began with at least 20 standard stimuli. The order of blocks was counterbalanced across participants.

While participants watched a movie or show with subtitles, stimuli were presented via headphones (HD 280 Pro, Sennheiser). Subjects were allowed a break after each block. The experiment lasted approximately 1.5 h.

Data acquisition and analysis

EEG was continuously recorded from 34 active Ag/AgCl electrode positions (actiCAP, Brain Products). The sampling rate was 1000 Hz, and data were filtered online from 0.1 to 1000 Hz.[6] FCz served as the online reference and AFz as the ground. Interelectrode impedances were kept below 25 kΩ. Subjects were asked to sit still and avoid excessive eye movements.

Offline data for each participant were re-referenced to the average of both mastoids and bandpass-filtered at 0.5–30 Hz. The data were segmented into epochs from 100 ms before to 600 ms after the onset of each sound. The first series of standards in each block, the first deviant in each block, and the first standard after each deviant were excluded from further analyses. Epochs were baseline-corrected using the pre-stimulus interval. Epochs in which the voltage at any channel exceeded ±75 μV were removed from further analysis. For each

participant, at least 60 deviant trials per condition were retained. Within each block type (FS and HS blocks), the MMN was calculated by subtracting the average ERP response to each standard from the average ERP response to the same phoneme when it was used as a deviant (e.g. standard-/f/ from one FS block was subtracted from deviant-[f] in the other FS block).

Statistical analysis of MMN amplitudes was conducted over the whole head from 100 to 300 ms (the time window in which MMN is expected to appear) using spatiotemporal clustering (Maris & Oostenveld, 2007) to correct for multiple comparisons. This method identifies clusters of spatially and temporally adjacent datapoints that meet an arbitrary threshhold of significance ($p = .05$ in the present analysis) and then evaluates the significance of the clusters using a nonparametric permutation statistic. This method allows for testing main effects and interactions over a broad temporal and spatial window, without having to choose a specific time window for each peak, and to then draw conclusions about the temporal and spatial distribution of the data contributing to any significant effects. For the purpose of plotting topographic maps, we also chose an MMN time window by separately identifying the peak latencies of the MMNs (at electrode Fz) for the FS blocks and for the HS blocks, and averaging the samples from 25 ms before the earlier peak to 25 ms after the later peak; the resulting time window was 173–248 ms.

Results

MMN difference waves at electrode Fz, as well as topographic maps for the 173–248 ms time window comprising both MMN peaks, are shown in Figure 3. In the FS blocks (left side) a strong asymmetry is evident, suggesting that a larger MMN was elicited when /f/ was standard and [s] was deviant than vice versa. This asymmetry pattern is weaker or nonexistent in the HS blocks (center). Statistical analysis confirmed these visual observations.

The CONTRAST × DEVIANTSEGMENT interaction effect yielded a significant effect ($p = .046$), which was based on a cluster extending from 163 to 265 ms and including the channels F3, Fz, F4, FC5, FC1, FC2, C3, Cz, C4, T8, CP5, CP1, P7, P3, Pz, PO9, O1, and Oz. Follow-up pairwise tests were conducted on the averages of all datapoints in this cluster for each condition. For the FS-blocks, the asymmetry between the two difference waves was significant ($t(23) = -4.143$, 95% CI = $-1.73 \ldots -0.58$, $p < .001$); for the HS-blocks, the asymmetry between the two difference waves was not significant ($t(23) = 0.26$, 95% CI = $-0.46 \ldots 0.60$, $p = .800$). In addition to the
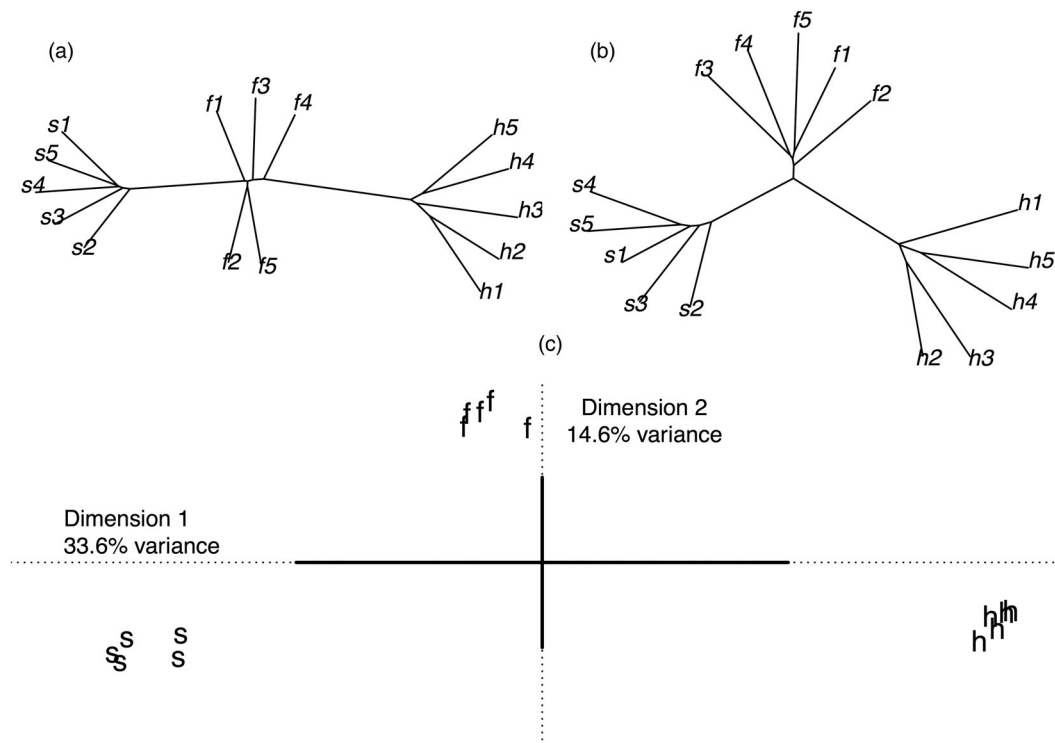
Figure 2. (a) Cluster analysis based on MFCC of the stimuli. (b) Cluster analysis based on delta coefficients of the stimuli. Both cluster analyses used the Neighbor-Joining Tree algorithm (Mielke, 2012; Saito and Nei, 1987). (c) Stimuli projected onto the first two dimensions of the compromise between the two distance matrices (Abdi et al., 2012). The first dimension clearly separates the segments [s], [f] and [h] along its axis in three small clusters, whereas the second dimension further separates the [s] and [h] clusters from the [f] cluster.

crucial interaction, the main effect of Contrast was also significant ($p = .002$), driven by a cluster from 202 to 297 ms, including all channels, indicating that the FS-blocks elicited more negative MMNs on average than the HS-blocks ($t(23) = -6.11$, 95% CI $= -1.74 \ldots -0.88$, $p < .001$); and the main effect of DEVIANTSEGMENT elicited a significant effect ($p = .002$) based on a cluster from 188 to 301 ms, including channels F7, F3, Fz, F4, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, CP5, CP1, CP2, CP6, P3, Pz, and P4, indicating that more negative MMNs were elicited when [s] was the deviant and /f/ or /h/ the standard than vice versa ($t(23) = -4.24$, 95% CI $= -1.68 \ldots -0.58$, $p < .001$).

For the sake of comparison with previous MMN studies, we also analysed the data using a repeated measures ANOVA over the pre-defined MMN time windows (see Data acquisition and analysis) on Fz, the channel where MMN is often strongest (e.g. Näätänen, 2001; Näätänen et al., 2007). The same pattern of results is also found if different MMN peak time windows are picked for each contrast or each condition. In this analysis, the ANOVA revealed a significant main effect of CONTRAST ($F (1, 23) = 7.50$, $p = .012$), indicating larger MMNs on average in the FS blocks than the HS blocks, and a significant main effect of DEVIANTSEGMENT ($F (1, 23) = 6.94$, $p = .015$), indicating smaller (less

negative) MMNs on average when /s/ was the standard (and [f] or [h] the deviant) than when /f/ or /h/ was the standard (and [s] the deviant). Crucially, there was also a marginal CONTRAST × DEVIANTSEGMENT interaction ($F (1, 23) = 3.53$, $p = .073$). The asymmetry in MMN amplitudes (i.e. the simple effect of DEVIANTSEGMENT) was significant in the FS blocks, where deviant-[s] with standard-/f/ elicited a more negative MMN than deviant-[f] with standard-/s/ ($t(23) = -3.31$, 95% CI $= -2.25 \ldots -0.52$, $p = .003$), but was not significant in the HS blocks ($t(23) = -0.37$, 95% CI $= -1.08 \ldots 0.75$, $p = .714$).

Discussion

The present study tested the predictions of FUL, a model of speech recognition, regarding the cognitive processing of fricatives in isolation. Using an oddball paradigm to test the neural response to the contrast between /f/ (which is fully specified for place of articulation) and /s/ (which is underspecified), we observed that, as predicted by FUL, the MMN response to deviants was attenuated in blocks with underspecified (/s/) standards compared to blocks with fully specified (/f/) standards. This result confirms and expands previous results obtained with stop consonants and vowels (Cornell et al., 2011, 2013; Scharinger & Lahiri, 2010; Scharinger, Lahiri, & Eulitz, 2010;
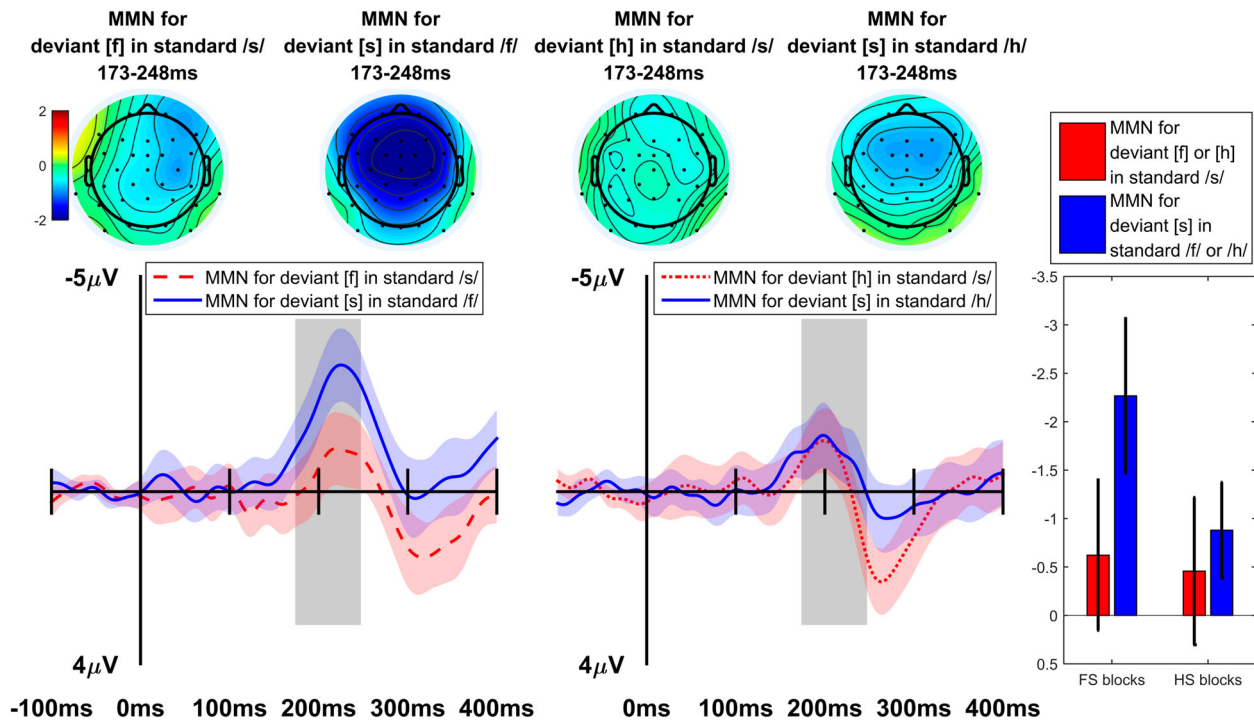
Figure 3. Difference waves (at Fz) consisting of the average of deviant trials for a given segment minus the average of standard trials for the same segment; topographic maps and bar plot of MMN mean amplitude during 173–248 ms time window. Both error bars on bar plot and width of ribbon represent ±2 standard errors.

Walter & Hacquard, 2004), while also demonstrating that presentation of single fricatives (and presumably sonorants) in isolation is a viable method of testing FUL's predictions, since they can be presented without any phonological context that could potentially influence the results (see Table 1 for unigram frequencies), as suggested by Bonte et al. (2005).

Furthermore, the present study used ERP data to elucidate the long-term memory representation of /h/ in terms of its featural content. When contrasting the fricative underspecified for place of articulation (/s/) against /h/, we did not observe a modulation of MMN amplitude, which would be expected if /h/ were specified for a place of articulation (e.g. [Radical]). This result provides new evidence that informs an existing theoretical debate about the featural specification of /h/. If the results of this study can be used as a diagnosis for featural content, /h/ appears to have no underlying [Radical] place of articulation on par with the [Labial] feature of /f/. With very similar symmetric MMN results between /s/ and /h/, we conclude that whatever the long-term featural representation of these two segments may be, they do not seem to clash. Thus, /h/ patterns with underspecified /s/ rather than specified /f/.

In what follows, we first discuss the functional significance of the asymmetry observed in the contrast between /f/ and /s/, and then lack of asymmetry in the contrast between /h/ and /s/ and its consequences for phonological theory.

The locus of the asymmetry between /f/ and /s/

We have attributed the asymmetrical MMN for the contrast between /f/ and /s/ as based on the underspecification of the [Coronal] feature when /s/ serves as a standard, as per FUL. While Gow (2001, 2002, 2003) argues that underspecification is not strictly necessary for place assimilation of oral and nasal stops, we argue that it may here be found in fricatives which do not alternate with one another. In this section we further address potential alternative explanations for why an asymmetrical MMN may arise.

As noted in the Methods, we controlled the intensity of the stimuli, as differences in intensity may modulate early ERPs regardless of linguistic factors. However, /f/ in English typically has a lower intensity than /s/ (McMurray & Jongman, 2011), such that intensity is likely a useful cue for the perception of this contrast, and normalising the intensity of the stimuli has neutralised this cue. Specifically, since we normalised to a relatively high volume, it is possible that we made the [f] tokens more /s/-like – that is, it is possible that our [f] tokens were somewhat atypical members of the /f/ category, whereas [s] tokens were relatively good, prototypical tokens.[7] Behavioural studies on

the Perceptual Magnet Effect have suggested that discrimination from a less prototypical to a more prototypical token is easier than vice versa (Sussman & Lauckner-Morano, 1995) – although a recent study suggests that this effect only occurs in discrimination between very similar tokens within a single category, but not between pairs of relatively different tokens (Masapollo, Polka, & Ménard, 2015). While this is a potential concern, previous electrophysiological evidence suggests that that this could not account for the pattern of results we observed, and in fact would predict the opposite asymmetry. Ikeda, Hayashi, Hashimoto, Otomo, and Kanno (2002) contrasted typical and atypical tokens of Japanese /e/, and found that the MMN was *smaller* when the standard was a poor token of /e/ and the deviant was a good token, and larger in the opposite case. In the present study, on the other hand, the MMN was smaller when the standard was [s] tokens, which arguably were more typical exemplars of their category. While there are differences between these studies (for instance, Ikeda et al. (2002) examined typical and atypical tokens of the same category, whereas the present study arguably examined typical tokens of one category and atypical tokens of another category), a category-typicality account seems less consistent with the present results than an underspecification account.

A related concern is that /f/, as a non-sibilant fricative, is intrinsically difficult to identify and to discriminate from other fricatives, and the frication alone provides few useful cues for perception (Jongman, Wayland, & Wong, 2000; McMurray & Jongman, 2011). Again, however, previous MMN literature would lead us to expect an attenuated MMN for a contrast between a hard-to-identify standard (/f/) and an easy-to-identify deviant (/s/), which is the opposite of what we observed.

Another acoustic concern is that /f/ typically has a more diffuse spectrum than /s/: in other words, at any given frequency band, /f/ tends to have less energy than /s/. This means that hearing a deviant /f/ token in a stream of /s/ tokens could amount to detecting a decrease in energy or an absence of information, whereas hearing a deviant /s/ token in a stream of /f/ tokens would amount to detecting an increase of energy or an addition of information.[8] There is indeed ample evidence that a deviant which is missing an acoustic feature present in the standards elicits an attenuated MMN (or no MMN at all) compared to a deviant which has an additional acoustic feature not present in the standards (Nordby, Hammerborg, Roth, & Hugdahl, 1994; Sabri & Campbell, 2000; Timm, Weise, Grimm, & Schröger, 2011; see also Czigler, Sulykos, & Kecskés-Kovács, 2014, for a corresponding effect in the visual counterpart of the MMN). While these studies all used non-linguistic stimuli, similar asymmetries based on

feature salience have been seen for linguistic stimuli in behavioural studies (Nielsen, 2011; Nielsen & Scarborough, 2015). Thus, assuming that having more energy in certain frequency bands is similar to having an additional feature (in the studies cited above, the "additional" or "missing" feature was typically white noise or sine tone overlain over another non-linguistic sound, or a transient click or frequency modulation within a non-linguistic sound), then this effect could also account for the asymmetry we observed between /f/ and /s/. On the other hand, at some times in some of our stimuli /f/ has more energy than /s/ across all frequency bands (perhaps because of the amplitude normalisation), which challenges this account. It is difficult to compare previous findings on asymmetries based on acoustic features, which are based on the addition of a qualitatively different acoustic feature, to the present stimuli, which if anything involve more or less spectral energy rather than the addition of qualitatively different energy; therefore, whether this effect could account for the present results remains an open question.

Another factor that can cause asymmetries in both MMN and behaviour is the focality or peripherality of the phonemes in question (Cowan & Morse, 1986; Polka & Bohn, 2003, 2011; Schwartz, Abry, Boe, Ménard, & Vallée, 2005). For instance, when contrasting a corner vowel like /i/, which is articulated near the periphery of the vowel space (i.e. the tongue body is as high and as far forward as it can be without producing a fricative), against a non-corner vowel like /e/, discrimination is easier (and MMNs larger) when the non-corner vowel serves as the standard and the corner vowel serves as the deviant. This effect is likely based on acoustics and on the architecture of the human auditory system, as it is observable even in non-native contrasts, but not in birds and cats (Polka & Bohn, 2003, 2011). As this line of research is based on vowel perception, however, it is not clear to us whether it can account for these results with fricatives. Specifically, more "focal" vowels have been defined as those in which consecutive formants converge (Schwartz et al., 2005; Polka & Bohn, 2011); fricatives, however, do not have multiple clear formants. While /s/ does have a more focal (less diffuse) spectrum than /f/ (although in our stimulus set /s/ still had a rather diffuse spectrum, with a rising shape but no clear peak), it is an open question whether the effect of formant convergence in vowels would also extend to an effect of spectral focality in consonants.

Ultimately, because the feature [CORONAL] is often argued to be universally underspecified (cf. Avery & Rice, 1989) and thus predicts asymmetries in the same direction across languages, it is difficult to rule out non-linguistic factors such as acoustic or neurobiological

mechanisms when these also predict asymmetries in the same direction. Further work is underway to test features which are argued to not be universal, which should demonstrate whether or not the phonological system of any given language is driving these effects (e.g. if the same set of stimuli produces asymmetries in the opposite directions for two participant groups with different language backgrounds), or some mechanism related to audition in general.

The phonological representation of /h/

We interpret the symmetric pattern of results for the /s/ and /h/ comparisons to indicate the features of /s/ and /h/ do not clash. On the basis of our data alone, it is not possible to determine whether /h/ has a phonetic or surface [RADICAL] place of articulation in the same way as /s/ has a [CORONAL] place of articulation: /s/ and /h/ could be similarly underspecified. Since a phonetic [RADICAL] feature may not clash with the underspecified place of /s/, the same *nomismatch* pattern is predicted for a phonetically specified but phonologically underspecified [RADICAL] [h] and a placeless /h/. One reviewer points out that /h/ has not always been classified as a fricative but rather as a glide. FUL suggests [SONORANT] and [OBSTRUENT] features ought to clash (Lahiri & Reetz, 2010), and this warrents further study as well. Nonetheless, there are theoretical reasons to expect that /h/ and /s/ must differ in some feature in order to be distinct speech sounds.

From FUL's perspective, it may be challenging for [h] to be specified at a phonetic level for a [RADICAL] place feature while being left underspecified for place of articulation at the phonological level. This is because FUL should not tolerate two default places of articulation without the addition of some mechanism able to distinguish underspecified /s/ from underspecified /h/.[9] The system would have no reason to specifiy underspecified segments as [RADICAL] and others as [CORONAL] in the absence of other features distinguishing the two segments. Thus, considering [RADICAL] to be a different sort of feature (a larynx or glottalic state feature, rather than a place of articulation) may make more sense. This is exactly the proposal of Goldsmith (1981)
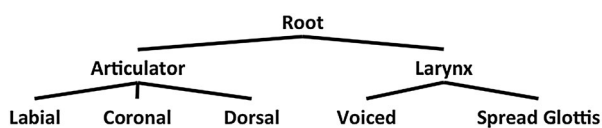

Figure 4. Feature geometric representation of oral and laryngeal articulators. The oral and laryngeal articulators may be represented distinctly, as in notes on a feature geometry tree.

wherein debucalization in some dialects of Spanish (i.e. /s/ to /h/ in coda position) is a result of the place node in a hierarchically organised set of features being completely lost. In the same vein, Davis and Cho (2003) suggests that in American English, /h/ is related to aspiration (i.e. [SPREAD GLOTTIS]). FUL already incorporates some aspects of feature-geometric theories as tongue height (i.e. [HIGH] and [LOW] are distinct from articulator or place of articulation) and since laryngeal gestures (i.e. voicing and aspiration) coexist with other consonants, the larynx may warrant the same treatment as an articulator distinct from the lips and tongue. In fact, FUL already incorporates such a division (Lahiri & Reetz, 2010). Crucially, [RADICAL] is not an articulator feature used for /h/ in English. In terms of feature geometry, the oral articulators and larynx are seen as distinct nodes in a hierarchy of features, non-comprehensively sketched in Figure 4, a proposal which FUL already posits that would be able to capture the ERP results observed in our comparison of /s/ and /h/.

If the interpretation that features clash within – but not across – articulatory nodes is correct, other nodes should likewise be orthogonal and invisible to a clash of features. FUL hypothesises a tongue height feature which may similarly be amenable to this sort of investigation with the English interdental and alveopalatal fricatives /T/ and /S/, which differ from /s/ only by [HIGH] and [LOW] features (Lahiri & Reetz, 2010).[10] Likewise, further investigation of the uvular and pharyngeal segments within FUL may illuminate the division between the oral articulators, tongue root, and larynx.

Indeed, if boundaries between phonemes are porous – such that an underspecified segment can be activated by multiple phonetic inputs – our measures of phonological neighbourhood density could benefit by incorporating feature-based distinctions, as they seem to provide a more precise metric of form similarity for lexical representations. Finally, these results are particularly challenging to theories of lexical representations according to which the latter are grounded purely within sensory (acoustic) memory, like some episodic theories (Goldinger, 1998). We must leave the investigation of how exemplar models can account for these asymmetries for future research.

If acoustic distance was at the basis of the MMN results described in the study, then the comparison of /s/ and /h/ in the HS blocks should have yielded the largest MMN responses. Contrary to this prediction, the largest MMN responses were obtained when comparing the acoustically more similar segments /s/ and /f/. Furthermore, if indeed the acoustic properties are driving these results, native language should have no bearing on these results. For a feature, such as voicing, which

is argued to parametrically vary between languages, we expect a different pattern for speakers of those languages. Work is underway to test this possibility. Likewise, frequency alone cannot account for the data pattern: in both measures, /f/ and /h/ had similar frequencies and /s/ was the outlier, in which case a frequency-based account would expect to see a different pattern for /s/ than for the other segments, whereas we observed a different pattern for /f/ than for the other segments.

Finally, we acknowledge an alternative interpretation for the lack of asymmetry between /h/ and /s/. In isolation, /h/ is phonetically realised as a voiceless vowel, and this was also the case in our stimuli (for instance, unlike /f/ tokens, /h/ tokens had some formant structure). In this case, participants may not have perceived the token as /h/ but may rather have perceived it as a vowel, in which case they would extract a [Vocalic] feature while hearing it, and this feature in turn would clash with the [Consonantal] feature of /s/. Since neither of these features is generally assumed to be underspecified, and indeed FUL (Lahiri & Reetz, 2010) suggests these two features must clash (though Avery & Idsardi, 2001, and references therein, suggest they need not necessarily clash), this would lead us to expect a symmetrical mismatch effect between the two; furthermore, this symmetrical mismatch effect may overshadow or supercede any potential asymmetrical effects based on place of articulation. While it is not possible to rule out this account on the basis of the present dataset, and it may be an inherent limitation of testing /h/ fricatives in isolation, we do note that this would not change the functional interpretation of the asymmetry between /f/ and /s/ discussed above. Furthermore, treating the /h/ tokens as voiceless vowels only accounts for our pattern of results (non-asymmetry) if we adopt several additional assumptions, namely, (1) that features are organised such that there is no underspecification of vocalic or consonantal features (this is indeed a common assumption, but not a universal one); and (2) that a symmetric feature clash between vocalic and consonantal features would override an asymmetrical place of articulation feature clash. This interpretation also conflicts with the observation that a larger MMN should be elicited from dissimilar stimuli (Sams et al., 1985).

Conclusion

We tested the predictions of a non-redundant, articulatory feature-based model (FUL) for the representational format of fricative categories /s/ and /f/ at the lexical level in an MMN paradigm. The results confirmed and expanded on existing findings suggesting that the feature [Coronal] is not used for lexical storage, being instead inserted as a default value. Moreover, we also compared /s/ and /h/ to test whether /h/ may be specified for a radical articulation on par with other places of articulation such as [Labial]. The segment /h/ did not pattern with /f/, suggesting that any laryngeal features of /h/ are not place features on a par with the [Labial] feature of /f/. Rather, the neurophysiological evidence suggests that /h/ either lacks a place of articulation altogether or, more plausibly, following the proposals of well-established feature geometry models Clements (1985); Sagey (1986); Goldsmith (1981); McCarthy (1988), the laryngeal features present in /h/ do no clash with the oral articulators of present in /s/.

Notes

1. Slashes / / represent long-term or phonological representations, while square brackets [ ] are used to denote short-term or phonetic representations. Features are indicated in small caps within square brackets.
2. See Haspelmath (2006) for a recent description, though here the notion of Jacobson and Trubetskoy is the most pertinent. Other theories may derive similar or better predictions with privative features or some other mechanism.
3. Other factors that can elicit asymmetries in MMN are addressed in the Discussion.
4. A reviewer points out that Lahiri and Reetz (2010) are not particularly reliant on this suggested representation for laryngeal sounds.
5. Or, as a reviewer points out, manner features like [Fricative].
6. Because of an error in amplifier setup, the online low-pass filter was set to 1000 Hz, which is above the Nyquist frequency for sampling at 1000 Hz. This means that the recorded signals may be aliased: signals above 500 Hz would not be correctly represented in the data (this is not a concern for the present study since such frequencies are beyond what is of interest in language processing, and will have been removed anyway through our offline low-pass filter) and, more importantly, spurious low-frequency signals may be introduced in the data (although note that we also used an offline high-pass filter). While this is indeed a concern, and it would be valuable to have the results replicated in future studies with a different setup, we also note that it is unlikely that this would have caused spurious between-condition differences in our results, because if any aliasing effects were not corrected by offline filters they would at least have happened throughout the whole dataset.
7. We thank the editor for pointing out this possibility.
8. We thank the editor for pointing out this possibility.
9. A reviewer points out this type of mechanism (i.e. primary vs. secondary place of articulation) may be exactly the type of mechanism which is independently needed to distinguish a labialized segment from a coarticulated segment (e.g. /gʷ/ from /g͡b/)

10. Other feature systems may use different features to distinguish these (e.g. Avery & Idsardi, 2001).

ORCID

Kevin Schluter http://orcid.org/0000-0003-2721-9400
Stephen Politzer-Ahles http://orcid.org/0000-0002-5474-7930
Diogo Almeida http://orcid.org/0000-0003-4674-8092

References

Abdi, H., Williams, L. J., Valentin, D., & Bennani-Dosse, M. (2012). STATIS and DISTATIS: Optimum multitable principal component analysis and three way metric multidimensional scaling. *Wiley Interdisciplinary Reviews: Computational Statistics*, *4*(2), 124–167. doi:10.1002/wics.198

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology*, *5*(2), 183–207. doi:10.1017/S0952675700002268

Archangeli, D. B. (1984). *Underspecification in Yawelmani phonology and morphology* (PhD thesis). Massachusetts Institute of Technology.

Avery, P., & Idsardi, W. (2001). Laryngeal dimensions, completion and enhancement. In T. A. Hall (Ed.), *Distinctive feature theory* (pp. 41–70). New York: Mouton de Gruyter.

Avery, P., & Rice, K. (1989). Segment structure and coronal underspecification. *Phonology*, *6*(2), 179–200. doi:10.1017/S0952675700001007

Baković, E. (2014). Phonemes, segments and features. *Language, Cognition and Neuroscience*, *29*(1), 21–23. doi:10.1080/01690965.2013.848992

Bonte, M. L., Mitterer, H., Zellagui, N., Poelmans, H., & Blomert, L. (2005). Auditory cortical tuning to statistical regularities in phonology. *Clinical Neurophysiology*, *116*(12), 2765–2774. doi:10.1016/j.clinph.2005.08.012

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook*, 225–252. doi:10.1017/S0952675700000440

Cornell, S. A., Lahiri, A., & Eulitz, C. (2011). "What you encode is not necessarily what you store": Evidence for sparse feature representations from mismatch negativity. *Brain Research*, *1394*, 79–89. doi:10.1037/a0030862

Cornell, S. A., Lahiri, A., & Eulitz, C. (2013). Inequality across consonantal contrasts in speech perception: Evidence from mismatch negativity. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 757. doi:10.1037/a0030862

Cowan, N., & Morse, P. (1986). The use of auditory and phonetic memory in vowel discrimination. *Journal of the Acoustical Society of America*, *79*, 500–507. doi:10.1121/1.393537

Czigler, I., Sulykos, I., & Kecskés-Kovács, K. (2014). Asymmetry of automatic change detection shown by the visual mismatch negativity: An additional feature is identified faster than missing features. *Cognitive, Affective, and Behavioral Neuroscience*, *14*, 278–285. doi:10.3758/s13415-013-0193-3

Davis, S., & Cho, M.-H. (2003). The distribution of aspirated stops and /h/ in American English and Korean: An alignment approach with typological implications. *Linguistics*, *41*(4), 607–652. doi:10.1515/ling.2003.020

Eulitz, C., & Lahiri, A. (2004). Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive Neuroscience*, *16*(4), 577–583. doi:10.1162/089892904323057308

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251. doi:10.1037/0033-295X.105.2.251

Goldsmith, J. (1981). Subsegmentals in Spanish phonology: An autosegmental approach. In W. Cressey & D. Napoli (Eds.), *Linguistic symposium on romance languages* (Vol. 9, pp. 1–16). Washington: Georgetown University Press.

Gow, D. W. J. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*, *45*, 133–159. doi:10.1006/jmla.2000.2764

Gow, D. W. J. (2002). Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance*, *28*(1), 163–179. doi:10.1037/0096-1523.28.1.163

Gow, D. W. J. (2003). Feature parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics*, *65*(4), 575–590. doi:10.3758/BF03194584

Halle, M. (1959). *The sound pattern of Russian*. Berlin: Mouton.

Haspelmath, M. (2006). Against markedness (and what to replace it with). *Journal of Linguistics*, *42*, 25–70. doi:10.1017/S0022226705003683

Hestvik, A., & Durvasula, K. (2016). Neurobiological evidence for voicing underspecification in English. *Brain and Language*, *152*, 28–43. doi:10.1016/j.bandl.2015.10.007

Ikeda, K., Hayashi, A., Hashimoto, S., Otomo, K., & Kanno, A. (2002). Asymmetrical mismatch negativity in humans as determined by phonetic but not physical difference. *Neuroscience Letters*, *321*(3), 133–136. doi:10.1017/S0952675710000102

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America*, *108*, 1252–1263. doi:10.1121/1.1288413

Kazanina, N., Phillips, C., & Idsardi, W. (2006). The influence of meaning on the perception of speech sounds. *Proceedings of the National Academy of Sciences*, *103*(30), 11381–11386. doi:10.1073/pnas.0604821103

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. *Laboratory Phonology*, *7*, 637–675.

Lahiri, A., & Reetz, H. (2010). Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics*, 38(1), 44–59. doi:10.1016/j.wocn.2010.01.002

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74(6), 431–461. doi:10.1037/h0020279

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. doi:10.1037/h0020279

Masapollo, M., Polka, L., & Ménard, L. (2015). Asymmetries in vowel perception: Effects of formant convergence and category "goodness". *Journal of the Acoustical Society of America*, 137, 2385. doi:10.1121/1.4920678

McCarthy, J. J. (1988). Feature geometry and dependency: A review. *Phonetica*, 45, 84–108. doi:10.1159/000261820

McCarthy, J. J., & Taub, A. (1992). Review of C. Paradis & J.-F. Prunet (Eds.) (1991). The special status of coronal: Internal and external evidence. *Phonology*, 9, 363–370. doi:10.1017/S0952675700001664

McMurray, B., & Jongman, A. (2011). What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations. *Psychological Review*, 118, 219–246. doi:10.1037/a0022325

Mester, A., & Itô, J. (1989). Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language*, 65(2), 258–293. doi:10.2307/415333

Mielke, J. (2012). A phonetically based metric of sound similarity. *Lingua*, 122, 145–163. doi:10.2307/415333

Mitterer, H. (2011). The mental lexicon is fully specified: Evidence from eye-tracking. *Journal of Experimental Psychology: Human Perception and Performance*, 37(2), 496–513. doi:10.1037/a0020989

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, 38, 1–21. doi:0.1111/1469-8986.3810001

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, 118(12), 2544–2590. doi:10.1016/j.clinph.2007.04.026

Newman, R., Clouse, S., & Burnham, J. (2001). The perceptual consequences of within-talker variability in fricative production. *Journal of the Acoustical Society of America*, 109, 1181–1196. doi:10.1121/1.1348009

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, 39, 132–142. doi:10.1121/1.1348009

Nielsen, K., & Scarborough, R. (2015). *Perceptual asymmetry between greater and lesser vowel nasality and VOT*. In Proceedings of the 18th International Congress of Phonetic Sciences.

Nordby, H., Hammerborg, D., Roth, W., & Hugdahl, K. (1994). ERPs for infrequent omissions and inclusions of stimulus elements. *Psychophysiology*, 31, 544–552. doi:10.1111/j.1469-8986.1994.tb02347.x

Perkell, J. S., & Klatt, D. H. (Eds.). (2014). *Invariance and variability in speech processes*. New York: Psychology Press.

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., … Roberts, T. (2000). Auditory cortex accesses phonological categories: An MEG mismatch study. *Journal of Cognitive Neuroscience*, 12(6), 1038–1055. doi:10.1162/08989290051137567

Polka, L., & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication*, 41, 221–231. doi:10.1016/S0167-6393(02)00105-X

Polka, L., & Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, 39, 467–478. doi:10.1016/j.wocn.2010.08.007

Sabri, M., & Campbell, K. (2000). Mismatch negativity to inclusions and omissions of stimulus features. *NeuroReport*, 11, 1503–1507.

Sagey, E. C. (1986). *The representation of features and relations in non-linear phonology* (PhD thesis). Massachusetts Institute of Technology.

Saito, N., & Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), 406–425.

Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology/ Evoked Potentials Section*, 62(6), 437–448. doi:10.1016/0168-5597(85)90054-1

Scharinger, M., Bendixen, A., Trujillo-Barreto, N. J., & Obleser, J. (2012). A sparse neural code for some speech sounds but not for others. *PLoS one*, 7(7), e40953. doi:10.1371/journal.pone.0040953

Scharinger, M., & Lahiri, A. (2010). Height differences in English dialects: Consequences for processing and representation. *Language and Speech*, 53(2), 245–272. doi:10.1177/0023830909357154

Scharinger, M., Lahiri, A., & Eulitz, C. (2010). Mismatch negativity effects of alternating vowels in morphologically complex word forms. *Journal of Neurolinguistics*, 23(4), 383–399. doi:10.1016/j.jneuroling.2010.02.005

Scharinger, M., Monahan, P. J., & Idsardi, W. J. (2012b). Asymmetries in the processing of vowel height. *Journal of Speech, Language, and Hearing Research*, 55(3), 903–918. doi:10.1044/1092-4388(2011/11-0065)

Schwartz, J.-L., Abry, C., Boe, L.-J., Ménard, L., & Vallée, N. (2005). Asymmetries in vowel perception, in the context of the dispersion–focalisation theory. *Speech Communication*, 45, 425–434. doi:10.1044/1092-4388(2011/11-0065)

Sussman, J. E., & Lauckner-Morano, V. (1995). Further tests of the "perceptual magnet effect" in the perception of [i]: Identification and change/no-change discrimination. *Journal of the Acoustical Society of America*, 97, 539–552. doi:10.1121/1.413111

Tavabi, K., Elling, L., Dobel, C., Pantev, C., & Zwitserlood, P. (2009). Effects of place of articulation changes on auditory neural activity: A magnetoencephalography study. *PLoS one*, 4(2), e4452. doi:10.1371/journal.pone.0004452

Timm, J., Weise, A., Grimm, S., & Schröger, E. (2011). An asymmetry in the automatic detection of the presence or absence of a frequency modulation within a tone: a mismatch negativity study. *Frontiers in Psychology*, 2(189). doi:10.3389/fpsyg.2011.00189

Vaden, K., Halpin, H., & Hickok, G. (2009). Irvine phonotactic online dictionary, version 2.0. [Data file]. Retrieved from http://www.iphod.com

Walter, M. A., & Hacquard, V. (2004). *MEG evidence for phonological underspecification*. In 14th International Conference on Biomagnetism, pp. 292–293.