

Published in final edited form as:

Lingua. 2014 February 1; 140: 35–51. doi:10.1016/j.lingua.2013.11.007.

Sparseness of vowel category structure: Evidence from English dialect comparison

Mathias Scharinger^{a,b} and William J. Idsardi^b

^aMax-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

^bDepartment of Linguistics, University of Maryland, USA

Abstract

Current models of speech perception tend to emphasize either fine-grained acoustic properties or coarse-grained abstract characteristics of speech sounds. We argue for a particular kind of 'sparse' vowel representations and provide new evidence that these representations account for the successful access of the corresponding categories. In an auditory semantic priming experiment, American English listeners made lexical decisions on targets (e.g. *load*) preceded by semantically related primes (e.g. *pack*). Changes of the prime vowel that crossed a vowel-category boundary (e.g. *peck*) were not treated as a tolerable variation, as assessed by a lack of priming, although the phonetic categories of the two different vowels considerably overlap in American English. Compared to the outcome of the same experiment with New Zealand English listeners, where such prime variations *were* tolerated, our experiment supports the view that phonological representations are important in guiding the mapping process from the acoustic signal to an abstract mental representation. Our findings are discussed with regard to current models of speech perception and recent findings from brain imaging research.

Keywords

Speech perception; phonetic and phonological categories; abstract information; exemplars; vowel representations; semantic priming; lexical decision

1. Introduction

1.1. Vowel variance

English vowels show substantial variation in pronunciation across speakers. This can arise from many factors, most prominently gender, dialect, and social background (e.g. Hagiwara, 1997; Johnson, 1997; Lindblom, 1990; Thomas, 2001). Acoustically, this variation can be measured by means of the first and second formant values (F1 and F2), being the two most salient acoustic cues for vowel identification and categorization across languages (Flege et al., 1994; Ladefoged, 2001; Lindblom and Studdert-Kennedy, 1967; Pols et al., 1969; Stevens, 1998). Nevertheless, vowel categories display considerable overlap in their F1 and F2 values (e.g. front vowels [æ], [ɛ] and [i]; cf. Hillenbrand et al., 1995; Peterson and

© 2013 Elsevier B.V. All rights reserved.

Corresponding author: Mathias Scharinger (PhD), Max Planck Institute for Human Cognitive & Brain Sciences, Stephanstr. 1A - 04103 Leipzig, Germany, Tel.: +49 (0) 341-9940-2482, mscharinger@cbs.mpg.de.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Barney, 1952, even when F1 and F2 values are corrected for vocal tract size, e.g. through F3 normalization, Monahan and Idsardi 2010). This suggests that vowel categories are inherently fuzzy and such impressions are strengthened by the findings that vowel perception is less categorical than (stop) consonant perception (Pisoni, 1973; Schouten and van Hesse, 1992). Yet, listeners can distinguish vowels in close vicinity (such as [ɛ] and [æ]) with high accuracy (Hillenbrand et al., 1995), suggesting that at least on one level of processing (or in one kind of experimental task), phonetically detailed information is available. We argue that despite these facts, it is feasible to maintain an approach of *abstract* vowel representations where representations are not only based on discrete features, but also lack some specific feature specifications (underspecification). While the results of the experiment reported here suggest that overlapping vowel categories in Standard American English (henceforth AE) do not cause perceptual ambiguities in a behavioral online task, the same experiment with New Zealand English (NZE) listeners showed a rather different pattern (Author and Author, 2010), providing evidence that phonological and categorical representations guide the mapping from acoustic signals to long-term representations of speech. We also relate our findings to a recent neurophysiological study that supports the assumption of abstract category representations (Author et al., 2012). The current as well as the previous results are discussed within three different representational approaches, assuming *exemplars* (Bybee, 2001; Bybee and Hopper, 2001; Pierrehumbert, 2002; Pierrehumbert, 2001), abstract *fully specified* (e.g. Chomsky and Halle, 1968) or abstract *underspecified* representations (Lahiri and Reetz, 2002, 2010). Overall, we argue that the approach defining vowel categories in terms of abstract phonological features, with some features lacking an underlying specification, provides the most parsimonious explanation of the current behavioral as well as neurophysiological data, as will be discussed in more detail below.

1.2. Units of representations

Research in speech perception has suggested a variety of perceptual units that may be active on different levels of processing during language comprehension. Theories have suggested (amongst others) *exemplars*, i.e. faithful and very detailed, episodic representations (Bybee, 2001; Pierrehumbert, 2002; on the general principles of episodic memory, see Hasselmo, 2012), and the *phonological feature*. The latter constitutes a sub-phonemic, contrastive unit, referring to specific acoustic properties concomitant to particular articulator configurations (Halle, 1983; Jakobson et al., 1952) that can be expressed in a binary (e.g. ±VOICE, cf. Chomsky and Halle, 1968) or privative manner (where voiced sounds are marked with [VOICE] and voiceless sounds entirely lack this feature, cf. Lombardi, 1996). Here, we will contrast models that assume exemplars with models that assume phonological features at different degrees of abstraction (fully specified versus underspecified).

1.3 Exemplar models

Exemplar models are characterized by exhaustive representation of phonetic detail along with non-linguistic (indexical) information, such as speaker or dialect information. While there are individual differences between specific exemplar approaches (Bybee, 2001; Goldinger, 1998; Nosofsky and Palmeri, 1997; Nosofsky and Zaki, 2002), most exemplar theories converge in assuming that representations consist of sets of remembered and very detailed tokens (single segments and/or words, see Figure 1), for which there is either a most typical member (prototype) or a category center (Bybee, 2001; Bybee and Hopper, 2001; Goldinger, 1996; Johnson, 2005; Pierrehumbert, 2002; Pierrehumbert, 2001; Pisoni, 1997; Thomas, 2004). As a consequence of these architectural assumptions, there are two crucial measures that should influence the speed with which a particular sound or word is mapped on a category representation. The first of these measures is frequency of occurrence. On the level of words (and similarly, on the level of single segments), representations with a higher

frequency of occurrence tend to have more exemplars and can thus be accessed faster and more efficiently. The second measure relates to the typicality of specific exemplars with regard to their category center or prototype, and can be assessed by the Euclidean distance in a multidimensional acoustic space. The assumption is that a novel stimulus can access a particular representation faster if the acoustic distance of this stimulus to the category center is smaller. This assumption is particularly relevant for simple exemplar models that are based on k-Nearest Neighbors (e.g. Dudani, 1976; Daelenmans and van den Bosch, 2005), where categorization of novel stimuli depends on their (acoustic) distance to (existing) category prototypes. Another important consequence of exemplar models is that variation is directly expressed in the structure of representations.

As illustrated in Figure 1 (left), overlaps of the mid vowel [ɛ] with either high or low vowels can be expressed by the inclusion of high and low vowels in the exemplar set. A low vowel can thus be tolerated as a variant for a mid vowel.

1.4 Fully specified feature models

Outside exemplar theory, one of the main motivations for a sub-phonemic unit is rooted in the effort to find invariant, possibly higher-order acoustic cues for units of speech perception (e.g. Kewley-Port, 1983; Stevens and Blumstein, 1978; Sussman et al., 1991). While it proved to be impossible to find simple invariant cues for phonemes due to the effects of coarticulation (Benguerel and McFadden, 1989; Fowler, 2005; Lively et al., 1994), assimilation (Cutler, 1997; Ernestus et al., 2006; Flemming, 2001; Gaskell and Marslen-Wilson, 1996; Gow, 2001; Jun, 2004; Marslen-Wilson et al., 1995; Snoeren et al., 2008) and further factors in natural speech such as deletions and reductions (Ernestus, 2000; Mitterer and Ernestus, 2006; Mitterer et al., 2008; Raymond et al., 2006; Zimmerer et al., 2008), the phonological feature appears to be a better candidate in this respect because it is oftentimes closer to acoustic or articulatory invariances (Blumstein and Stevens, 1981; Lahiri et al., 1984). The expression of a phonological contrast by means of abstract features has a long tradition in linguistic theory (Fant, 1960; Jakobson, 1939; Jakobson et al., 1952), although theories differ as to whether phonological features are predominantly determined by articulatory gestures (Browman and Goldstein, 1992) or by acoustic cues produced by these gestures (Diehl, 2000; Diehl et al., 1991; Stevens, 2002), or through a combination of articulatory and acoustic factors (Jakobson et al., 1952).

One of the motivations for fully specified features in underlying representations is formulated in Halle et al. (2000) and Calabrese (1995). In particular, Calabrese (1995) argues that full representations are necessary in order to distinguish between phonological rules that are sensitive to either contrastive, marked, or both, contrastive and marked features. Another motivation for fully specified representations is embedded in the architecture of Optimality Theory (Prince and Smolensky, 2004), particular with regard to Lexicon Optimization. Note however that some Optimality Theory approaches also embrace the notion of underspecification (Inkelas, 1995; Itô et al., 1995).

Feature models accounting for the representation of mid vowels can follow full specification approaches or underspecification approaches, as introduced in the next section. A full specification approach would stipulate a specific feature representation by means of binary values, expressing that a mid vowel is neither high nor low. Such representations may follow the featural assumptions of Chomsky and Halle (1968) and subsequent accounts that provide full feature specifications for mid vowels (illustrated in Figure 1, middle), usually [-high, -low]. The mapping of vowel variants onto mid vowel representations does not follow straightforwardly from the representational architecture and appears to require a licensing rule that would interpret low vowels as [-low] in certain (non-low) contexts, the reverse of segmental redundancy rules. From a processing perspective, such an additional mechanism

is certainly time-consuming such that a low vowel is accepted as variant for a mid vowel only after additional processing time. For this reason, we contrast a full specification approach with an underspecification approach illustrated below.

1.5 Underspecified model

Underspecification theory has a relatively long, albeit not undisputed tradition in Generative Phonology (McCarthy and Taub, 1992). Here, we consider a particular, psycholinguistically motivated approach to underspecification (the Featurally Underspecified Lexicon, FUL, Lahiri and Reetz, 2002, 2010), that bears some resemblance to Radical Underspecification (Archangeli, 1988), but also shows crucial differences. One of the most important differences is that FUL assumes privative features and the lack thereof in underspecified representations, while Radical Underspecification allows for negative feature specifications in the unmarked case (for a more detailed discussion, see Lahiri and Reetz, 2010 and Author, 2009). For instance, voiced plosives in FUL would be expressed by [VOICED] (not [+VOICED]), and voiceless plosives are represented without specification for their laryngeal feature (not as [–VOICED]).

FUL is thus based on representational underspecification that is motivated by various asymmetries, for example those found in assimilation processes (Paradis and Prunet, 1991). For instance, coronals such as [n] very often assimilate to the place of articulation of their following neighbors (e.g. lean bacon > *leam* bacon due to labial [b]) and are therefore assumed to be underspecified for place of articulation. They receive their place of articulation feature either by default, or by the spreading of features of neighboring, specified speech sounds. In contrast, non-coronals such as [m] hardly ever assimilate to the place of articulation of their neighbors (e.g. rum toffee > *run* toffee due to coronal [t] is hardly ever encountered) and are therefore assumed to be specified for place of articulation (here: LABIAL). In general, FUL assumes that contrastive features are represented only if they cannot be derived by redundancy rules, i.e. tries to make the most economic use of a (universal) set of features for a given language. For instance, in a vowel system with a three-way tongue height distinction between the positions high, mid, and low, mid vowels are articulated close to the resting vertical position for the tongue and do not need to have a tongue height specification. They are neither high nor low, and thus considered ‘underspecified’ for tongue height even though it is possible to define acoustic parameter values that correlate with mid vowels (viz. an intermediate value for F1).

FUL is not only a theoretical account of the nature of underlying representations, but also an approach of mapping acoustic information onto long-term memory representations of speech sounds. To that end, the mapping is based on a feature-by-feature comparison between information in the incoming signal and underlying representation within one feature dimension (e.g. tongue height). The success of activating a speech sound memory representation is dependent on a three-way result of *match*, *mismatch*, and *no mismatch*. A match occurs if the feature extracted from the signal directly corresponds to the feature in the sound’s underlying representation (e.g. [high] matches [high]). A mismatch occurs if the extracted feature is incompatible with the underlying feature specification. This is the case whenever feature specifications are opposed to each other (e.g. [high] mismatches with [low]). A no mismatch, finally, occurs if a feature is extracted from the signal, but there is no corresponding feature specification underlyingly, for example when [coronal] is recognized in the signal for an incoming [n] but matched against the lexical representation of /n/ without a place of articulation specification. By contrast, mid vowels are featurally coded without a specific representation for tongue height underlyingly, while it is still possible that the phonetic feature [MID] is extracted from the acoustic signal of a mid vowel. This leads to very specific predictions for the activation of high, mid, and low

vowels. First, high and low vowels always mismatch, i.e. high [ɪ] could never activate low /æ/. Second, both [æ] and [ɪ] should not be mismatches for mid /ɛ/ without a tongue height specification, i.e. both high and low vowels should similarly activate the mid vowel as it is stored without any height specification in memory (illustrated in Figure 1, right panel). Finally, the mid vowel itself does not match high or low vowels, if the tongue height feature [mid] is extracted and evaluated against [high] or [low].

FUL (with privative features and underspecification) therefore differs from approaches of full specifications (with binary features and no underspecification) with regard to theoretical points of view (e.g. expression of marked segments) as well as with regard to issues of processing (e.g. co-activation of mid vowels by high or low vowels). Together with Exemplar Approaches, all models briefly introduced above can account for vowel variation, and how vowel variants can access the correct vowel categories. They differ, however, with respect to processing speed and with respect to non-linguistic influences onto the mapping process. For testing differences in this mapping process, a cross-linguistic or cross-dialectal comparison is optimal, because it allows comparing the interpretation of identical acoustic vowel sounds across different listeners with substantially similar grammars in other respects. In this respect, it is important to note that vowel variation can be observed not only within a language (e.g. American English), but also between dialects or local varieties of a language (e.g. American English, AE vs. New Zealand English, NZE). The latter comparison is particularly crucial for mid vowel representations, since NZE shows a height shift of its front vowels compared to the AE counterparts while otherwise sharing much of the pan-dialectal English phonology. The NZE shift will be elucidated in greater detail in the next section.

2. Comparing English dialects

2.1. Changes in the New Zealand vowel system

Some English varieties show a particular kind of dialectal variation that is characterized by shifts of certain vowels across categorical and phonological boundaries (Labov et al., 2006). An example is provided by the short front vowels of NZE (Figure 2) that underwent a push-chain shift and a subsequent restructuring of the vowel space (Bauer, 1986; Maclagan and Hay, 2004).

Here, original low [æ] (e.g. in TRAP, as assessed from other dialects, see Wells, 1982) has moved to the mid position of [ɛ] (as in DRESS), while [ɛ] has moved to the position of high [ɪ] (as in KIT, see Gordon et al., 2004; Langstrof, 2006 for an account of the phenomenon). The latter vowel shifted towards the center of the vowel space and is alternatively transcribed as [ɨ] or [ə] (see Hawkins, 1976; Wells, 1982b for a more detailed discussion).

Figure 3 illustrates the pronunciation consequences for the nouns *pack*, *peck*, and *pick* in comparison to Standard American English (AE). The column ‘lexical set’ refers to a dialect-independent label of the respective vowels using a word representative of the class (Wells, 1982a). The NZE vowel shift thus resulted in a mid TRAP vowel and a high DRESS vowel, while the KIT vowel has become centralized.

While there is extensive work on acoustic and phonetic changes of the NZE front vowel shift (Hawkins, 1973; Bauer, 1986; Bell, 1997; Watson et al., 1998; Easton and Bauer, 2000; Watson et al., 2000; Langstrof, 2003; Gordon et al., 2004; Maclagan and Hay, 2004; Langstrof, 2006; Warren and Hay, 2006), only few studies provide dialect-internal diagnostics for phonological changes. However, Author and Author (2010) provided experimental evidence for the TRAP vowel to not only differ in its phonetic realization, but also in its phonological representation in comparison to its representation AE. Together with

longitudinal evidence (Maclagan and Gordon, 2004), demonstrating that the TRAP vowel was affected first by the push-chain, it seems safe to assume that the NZE TRAP vowel has indeed a different phonological representation than its AE counterpart. In terms of underspecification, we assume the NZE TRAP vowel to be underspecified for height. Note that on the basis of Lahiri & Reetz, the TRAP vowel (as front vowel) is also underspecified for its place of articulation. This assumption (universal underspecification of coronality for both vowels and consonants) would hold independently of the push-chain effects, i.e. the TRAP vowel is assumed to be underspecified for place of articulation in NZE and AE, independent of its tongue height.

To conclude, the crucial difference between the NZE and AE vowel systems is that AE has a three-way height contrasts for its front vowels (high, mid, low), while NZE has only a two-way height contrast (high, mid). The changed representations of formerly low vowels can be modeled with exemplar approaches, fully specified, and underspecified accounts, but only the latter framework allows for the most parsimonious explanation of the existing behavioral and neurophysiological data, as will be shown in the next section.

2.2. Experimental assessment of vowel representations

Compared to AE, the TRAP vowel in NZE is not low anymore, an observation that can be modeled by a shift of exemplar members towards more mid vowel tokens, i.e. a change in their numeric parameters without any further concomitant change in an abstract representation. That is, under such assumptions, the category has simply moved in phonetic space.

In feature-based approaches, this would either mean that the TRAP vowel is now specified for [-high, -low], or not specified at all, i.e. [—]. All models seem to converge in agreeing that the TRAP vowel in NZE would allow for more variation than the TRAP vowel in AE, as a result of mid vowels showing more variation than either high or low vowels and due to the lack of another contrasting low front vowel. However, the three approaches differ as to whether the processing of vowel variants that cross a category boundary would be accompanied by frequency, typicality, or processing time effects. To reiterate the assumptions of exemplar, fully specified and underspecified models, we argue that exemplar models would predict the processing of TRAP vowel variants to depend on the frequency of occurrence of the word containing this vowel (which substantially match the frequencies of TRAP vowels observed in AE due to the generally slower pace of lexical change between dialects), or on the acoustic distance of the variant to the category center of the TRAP vowel. Fully specified models, on the other hand, would presumably predict an effect of increased processing time, independently of the vowel's representation, while the underspecified model as introduced above claims that mid vowels can be accessed by high and low vowels alike, such that the NZE TRAP vowel should tolerate both high and low vowel variants because its featural representation is now underspecified in NZE but not in AE.

Author and Author (2010) examined these assumptions by employing an auditory semantic repetition priming experiment. This behavioral experimental technique has shown to reflect lexical organization and processing (Forster, 1999). Generally, priming refers to the observation that target stimuli can be accessed faster if they are preceded by prime stimuli that are related to them on a form-based or meaning-based level. Access latencies are measured by lexical decisions on the target words, commonly consisting of word/pseudo-word responses. In this respect, *pack* would speed up the access to the target *load* because of the shared meaning of these two nouns (Meyer and Schvaneveldt, 1971). Author and Author (2010) extended this paradigm in assuming that as long as nouns can be considered tolerable alternatives to *pack*, they would similarly prime the target *load*. In particular, they assumed

that NZE *pack* (and in general, all nouns with a TRAP stem vowel) can be accessed by either *peck* or *pick*, because the TRAP vowel in NZE has moved to a mid vowel position.

Note that for this assumption, it is not necessary to assume differences in phonological representations of the DRESS and KIT vowels. What is crucial, though, is the assumption phonetic feature extraction for the mapping of acoustic information onto long-term (abstract) memory representations. Supported by a large body of acoustic evidence (Hawkins, 1973; Bauer, 1986; Bell, 1997; Watson et al., 1998; Easton and Bauer, 2000; Watson et al., 2000; Langstrof, 2003; Gordon et al., 2004; Maclagan and Hay, 2004; Langstrof, 2006; Warren and Hay, 2006), the DRESS vowel in NZE can be considered [HIGH] with regard to tongue height, while the KIT vowel moved towards the back (Bauer, 1986, 1995; Gordon et al., 2004; Hay et al., 2006; Maclagan and Hay, 2004). We hypothesize that this backing is expressed by the feature [DORSAL] (in the notation of FUL for back vowels) that is extracted as acoustic-phonetic feature. We do not make claims about the phonological status of the DRESS and KIT vowel here. Such an account is clearly beyond the scope of this article and should be subject to future research. Note, however, that there exists some evidence regarding changes in the phonological awareness of these vowels (e.g. spelling changes: *flush* instead of *fish*). Thus, both in terms of tongue height and place of articulation, for NZE listeners, the DRESS and KIT vowel do not provide featural mismatches to the underspecified TRAP vowel (Author & Author, 2010).

In order to test this hypothesis, the authors measured the lexical decision times to targets semantically related to nouns containing the TRAP vowel (e.g. *load* in relation to *pack*, cf. Figure 3). Aside from conditions in which the TRAP vowel noun (e.g. *pack*) was used as prime for *load*, they also selected DRESS and KIT vowel primes (e.g. *peck*, *pick*). The three test conditions were compared to a control condition in which the target *load* was preceded by an unrelated noun (*bass*).

Exemplar models, fully specified and underspecified feature models would differ in their predictions regarding the modulations of lexical decision times on the targets across the different conditions. Within an exemplar framework, reaction times should correlate with word frequencies and with the acoustic distance between the prime vowel and the TRAP vowel. The latter prediction should be particularly relevant for simple exemplar models based on k-Nearest Neighbors (e.g. Dudani, 1976; Daelenmans and van den Bosch, 2005).

Fully specified models without further stipulations, on the other hand, would predict that neither *peck* nor *pick* could prime *load* as well as *pack* does because any such priming would have to depend on a specific rule that allowed the mapping of a high vowel onto a mid vowel, being specified for [-high, -low], an otherwise unmotivated innovation. The underspecification approach, as illustrated before, assumes that all three nouns (*pack*, *peck*, *pick*) can prime *load*, because neither of the DRESS and KIT vowels provides mismatching tongue height information for the underspecified TRAP vowel in NZE. Additionally, this model also predicts that response latencies for the target *load* do not correlate with frequency or acoustic distance of the prime vowels to the TRAP vowel.

Intriguingly, Author and Author (2010) found that all three prime types resulted in significantly faster lexical decision times on the respective targets compared to an unrelated control condition. They interpreted these findings as evidence for representational differences of the TRAP vowel between NZE and AE: The absence of the feature [low] in the TRAP words allowed for the partial mapping of a phonetically high vowel (as in NZE *peck*) onto the phonological TRAP category (cf. Figure 3), that is, NZE *peck* ([pɪk]) with [high] was a no-mismatch to the underspecified vowel in NZE *pack* (/pɛk/) in the terminology of Lahiri and Reetz (2002), and consequently the input *peck* facilitated the

activation of the memory representation of *pack*. That is, the successful mapping of *peck* to *pack* lead to facilitated lexical access of *pack* and therefore, to its semantic relatives (including the target *load*).

Exemplar models can only partially account for these findings. Note that the framework has been applied to NZE before (Warren and Hay, 2006; Warren et al., 2008) and may explain the overall priming pattern. As particularly illustrated in Thomas (2004), an exemplar approach can readily account for probability distributions of pronunciation differences regarding the TRAP vowel. In NZE, the exemplar distribution of words with this vowel has many more DRESS vowel realizations than TRAP vowel realizations, and perhaps even some KIT vowel alternatives. Therefore, if there is a high probability for certain pronunciation alternatives, these alternatives should cause priming of the respective exemplar distribution, which in turn should depend on the frequency of the corresponding word and the likelihood for alternative vowel realizations. In this regard, exemplar models differ from the assumption of the FUL model: in exemplar models, abstraction from statistically describable input representations may emerge a posteriori, as envisaged in e.g. Goldinger (2007) or Luce et al. (2003), i.e. bottom-up, while abstraction in Lahiri and Reetz (2002, 2010) is a model-based consequence of underlying representation from which differences in output distributions may emerge.

However, exemplar models would also predict additional interacting effects of word frequency and acoustic distance. Author and Author (2010) used frequency as a covariate in their statistical analyses and found no effect, thereby providing evidence that their pattern of results did not show a significant correlation with the frequency of the target stimuli for which lexical decisions were required. They also calculated correlations of lexical decision latencies and acoustic distance between the prime vowels and the TRAP vowel. Again, this correlation was not significant, suggesting that the priming pattern in fact reflected the nature of the underlying representation, as expressed in phonological features. It is important to note that the experiments in Author and Author (2010) were not designed to directly test frequency or phonetic similarity effects. Nevertheless, the lack of correlations between response latencies and either frequency or similarity measures suggests that the level of processing these latencies tap into is rather sensitive to the nature of underlying (concrete) representations. One might espouse an alternative interpretation of Author and Author's data according to which the more inclusive mapping to the TRAP vowel category in NZE would follow from the fact that this category has more varied exemplars¹. Such an account – reminiscent of an Exemplar approach – is not at odds with the FUL model, and we do not argue that this interpretation is incorrect. We argue, however, that FUL provides a theoretical account of *why* the trap vowel category has more varied exemplars, namely, by means of tongue height underspecification. This is not to say that underspecified speech sounds would always display more variation – a more detailed discussion of this empirical problem is beyond the scope of this article and must await future research, particularly for cases where fully specified sounds would show more variation.

Finally, a fully specified account could not explain the finding of Author and Author (2010), either, because there was no significant priming difference between the three prime type conditions. Fully specified accounts would predict significant differences in lexical decision times as soon as the prime variants crossed categorical boundaries, i.e. if *peck* precedes *load*.

In a second experiment, Author and Author (2010) presented the same NZE stimulus material to AE listeners whose TRAP vowel ought to be low. Crucially, the priming pattern was quite different from the one observed for the NZE listeners. While *pack* still primed

¹We thank an anonymous reviewer for this observation.

load, neither *peck* nor *pick* lead to significant priming. The authors argued that this pattern was attributable to the specification of the TRAP vowel in AE. Here, the specific vowel height feature [low] did not allow non-low vowel variants to be mapped onto its underlying representation. This rests on the assumption that AE listeners were still likely to extract a height feature (i.e. [LOW]) from the NZE trap vowel, in which case there was a full match between extracted feature and underlying representation. Furthermore, as for NZE listeners, Author and Author (2010) did not observe any statistically significant correlations between response latencies and either target word frequency or acoustic prime-target vowel distance. Considering both patterns of results, the best model to accommodate them appears to be the underspecification model as developed by Lahiri and Reetz (2002, 2010).

For a complete picture, two more experiments would be necessary. In the orthogonal combination of speaker/listener and AE/NZE, Author and Author (2010) tested NZE and AE listeners with stimuli from an NZE speaker and showed that the same acoustic input resulted in intriguingly different behavioral responses. These differences allowed for the interpretation that the TRAP vowel differs in its underlying representation between NZE and AE. Presenting AE listeners with the same stimulus material as used in Author and Author (2010) recorded by an AE speaker would eliminate the possibility that the priming pattern was a result of the listener not being familiar with the peculiar NZE pronunciation. Presenting NZE listeners with the stimulus material recorded by an AE speaker, on the other hand, would only serve the purpose of replicating the finding that NZE listeners are more tolerable to map vowel variants onto TRAP vowel representations. Furthermore, the latter experiment would be confounded by the fact that NZE listeners are highly familiarized with AE vowel pronunciations as a result of US media and movie exports to New Zealand, a fact that does not apply in the reverse case. Consequently, it is not clear what this experiment would add to the existing pattern of results. For these reasons, we decided to prioritize the within-dialect case in which AE listeners are exposed to the stimulus material recorded from an AE speaker. This experiment is described in more detail in the next section.

3. Auditory priming in American English with vowel variants

3.1 Motivation

In the present study, we intend to test whether the obtained priming pattern of Author and Author (2010) was an artifact of the phonetic material (the NZE stimuli) or indeed reflected the listeners' underlying representation of the TRAP vowel. For that purpose, we employed the same priming design as in the previous study, but used stimuli recorded from an AE speaker and presented them to AE listeners. If it is in fact the phonological vowel representation that accounted for the priming pattern in Author and Author (2010), and not the dialect of the speaker, we expect that *all* prime variants (containing both KIT and DRESS vowels) are not accepted as exemplars of the TRAP category for AE listeners and thus should not prime. Based on previous studies (Allen and Miller, 2001; Miller, 1995), we further expect a relatively clear extraction of phonetic category labels for the vowel exemplars presented. We also predict that the response latencies do not correlate with target word frequencies or acoustic distances between prime and target vowels. On the basis of the previous discussion, we expect the underspecification account to make the most precise predictions for the priming patterns.

3.2. Material

Sixteen triplets of English minimal pairs with stems containing the KIT, DRESS or TRAP vowel, were selected as primes (e.g. *pack*, *peck*, *pick*). This resulted in 48 monosyllabic test primes with a mean length of 3.3 segments. Corresponding targets were nouns semantically related to the prime with the TRAP vowel with a mean length of 3.2 segments (e.g. *load* for

pack). Targets were chosen on the basis of Webster's dictionary of synonyms (Gove, 1968) and were mostly (near)-synonyms of the TRAP vowel primes. Targets had no semantic relation to the other primes of the same triplet. Targets were also approximately matched to the frequency of their primes (48 per million [targets] vs. 49 per million [primes], based on COBUILD Spoken Word Frequency as determined from CELEX (Baayen et al., 1995).

Primes and targets were distributed over four experimental lists in a Latin Square design. This guaranteed that each participant heard a target only once while across participants, each target could be paired with four different prime types: (1) TRAP vowel nouns (e.g. *pack*); (2) DRESS vowel nouns (e.g. *peck*); (3) KIT vowel nouns (e.g. *pick*) and (4) unrelated nouns (e.g. *sense*). There were 16 critical item pairs in each participant list together with 52 filler pairs, 34 of which had pseudo-words at their second (target) position. This guaranteed that out of the total of 68 prime-target pairs, 34 required a word response on the target, and 34 a non-word response. Pseudo-words were derived from existing English words by altering 1–3 segments. They were phonotactically legal and assessed for their validity by a native English speaker. All item pairs were pseudo-randomized. The design of the auditory priming experiment is illustrated in Figure 2.

Stimuli were recorded from a native AE speaker with a clear three-way height distinction of the front vowel. The recording was done in a sound-attenuated room, using a Sony ECM88 microphone. Stimuli were directly recorded onto hard disk within the sound application PRAAT (Boersma and Weenink, 2011), with an amplitude resolution of 16 bits and a sampling rate of 44.1 kHz. Subsequently, the first two resonance frequencies (F1 and F2) of the test prime vowels were calculated on the basis of a Fourier analysis (FFT) that involved a 25 ms Hanning window, moved along the time dimensions in 5 ms steps. These analyses were carried out in PRAAT. Additionally, we calculated the Euclidean distance between each prime vowel variant (KIT, DRESS) and the TRAP vowel in the two-dimensional F2 x F1 space. Details about targets, control- and test-primes are provided in the appendix, with Appendix 1 displaying log-transformed surface frequencies for control primes and target words, and with Appendix 2 showing log-transformed surface frequencies of the test prime triplets, together with measures from the acoustic analyses.

3.3. Subjects and Procedure

68 students of the University of Maryland (52% females, mean age 20, SD=2.7) participated for class-credit and were randomly assigned one of the four experimental lists. They were tested individually and familiarized with the experimental task in a practice session with 10 prime-target pairs not occurring in the main experiment. In both tasks, experimental items were presented pair-wise, using the DMDX software (Forster and Forster, 2003), and participants had to provide a word/pseudo-word decision on the second member of each pair (the target). In order to match the attribution of button presses to the study by Author and Author (2010), right-handed subjects had to give word-responses by pressing J with their right index finger (pseudo-words: F), while left-handed subjects (N=2) were given the opposite instructions.

First (= prime) and second (= target) member of each experimental pair were separated by 250 ms, and subjects could respond within 2000 ms after target presentation. Reaction time measurement started at the onset of the target (see Figure 4).

The experiment, including a short briefing and the practice session, lasted about 15 minutes.

3.4. Results

Subjects showed acceptable performance on the targets (6.3% wrong responses, 0.8% time-outs), although 4 participants and 2 items had error-rates above 25% and were excluded from further analyses.

The dependent measures *accuracy* (correct vs. incorrect lexical decision) and *response times* (in particular the log-transformed response time values) were analyzed in Generalized Linear Mixed-Effect Models (Baayen et al., 2008; Pinheiro and Bates, 2000) with *subject* and *items* as random effects and *prime type* (TRAP, DRESS, KIT, control; treatment coding) as a fixed effect. For the accuracy analysis, we calculated a Mixed-Effect Logit Model (Agresti, 2002; Breslow and Clayton, 1993). Mixed-Effect Logit models provide a relatively new extension to Generalized Linear Logit Models. They allow for the inclusion of random effects and are less likely to yield spurious effects (Jaeger, 2008). In this model, we found an effect of *prime type* DRESS (Wald- $z = -2.30$, $p < 0.05$), reflecting higher error rates for the DRESS primes compared to the control condition (Table 1). The effect of prime type KIT showed a trend in the same direction (Wald- $z = -1.86$, $p = 0.06$), similarly reflecting higher error rates for the KIT primes compared to the control primes.

For the response time analyses, we additionally removed outliers with more than 2.5 SD from the mean (which amounted 8.5% of the data points). There was a main effect of *prime type* ($F(3,747) = 3.76$, $p < 0.05$), reflecting significant longer latencies in the control than in the TRAP vowel condition ($t = -2.20$, $p < 0.05$). The DRESS and KIT vowel condition did not differ significantly from the control condition ($ts < 1$, n.s.; cf. Table 2 and Figure 5). Note that the same analysis without outlier exclusion yielded very similar results (shorter latencies in the TRAP vowel condition than in the control condition [$t = -2.01$, $p < 0.05$]; similar latencies in the DRESS [$t = 0.4$, $p = 0.69$] and KIT vowel condition [$t = 0.48$, $p = 0.63$], compared to the control condition)

In order to assess potentially confounding and extraneous factors, we determined the best-fit mixed-model for response times through model comparison and criticism using a method parallel to that employed in multiple regression analyses (cf. Crawley, 2005) with the alternative fixed effects *phonemic length of prime*, *phonemic length of target*, *frequency of prime*, and *frequency of target* (COBUILD log frequencies from CELEX, Baayen et al., 1995, cf. Appendix 1 & 2). Note that the motivation of including frequency measures for both the prime and the target was based on the assumption that apart from the attested frequency effect on targets (Forster, 1999), lexical decision times could also be affected by frequency differences of the preceding primes, albeit in a different way. In general, we were interested in how far these measures (including prime and target frequency) would correlate with lexical decision times on the targets.

As a result of our model selection procedure, the best-fit model included the random effects *subject* and *item* and the fixed effects *phonemic length of target* and *frequency of target* in a fully factorial design. This model revealed no main effects of *phonemic length of target* or *frequency of target* ($ts < 2$, n.s.). Only the interaction *phonemic length of target* x *frequency of target* was significant ($t = 6.57$, $p < 0.05$). In particular, for low frequency targets, response times were faster with increasing target length.

We finally calculated a correlation analysis between response latencies and acoustic distances for a subset of the data. Acoustic distances were based on Euclidean distances between the prime variant vowels (KIT, DRESS) and the TRAP vowel. The rationale of this analysis is that exemplar models would predict a positive correlation between response latencies and distances, with longer latencies for vowel variants further away from the TRAP vowel. Such variants would be less probable exemplar tokens and should prolong

category access. Similar to previous findings, this correlation was not significant and actually showed a trend in the opposite direction ($r = -0.05$, $t = -1.07$, $p < 0.2$).

In sum, the results of this study showed

- a. less accurate lexical decisions for primes whose stem vowel deviated from the stem vowel of the semantically related prime (i.e. KIT or DRESS vowel instead of TRAP vowel)
- b. faster response latencies only in the TRAP vowel condition (i.e. for the semantically related primes)
- c. no significant correlations of latencies with other acoustic or lexical measures

These findings are discussed in more detail in the next section.

4. Discussion

4.1. Support for underspecified representations

The experiment reported here provides an important extension of the study by Author and Author (2010) and provides additional evidence for a particular kind of vowel representations that determine how vowel variants can be mapped onto underlying categories. As in the previous study, this experiment consisted of a modified auditory semantic priming paradigm, including – aside from semantically related prime-target pairs (e.g. *pack-load*) – variants of the primes that crossed a category boundary (e.g. *peck-load*). With this design, we wanted to test whether a deviance in the respective prime vowel would still yield significant target priming. We expected that this is only possible if the differing vowel is accepted as a variant of the vowel in the semantically related prime. This should result from a comparison of the variant prime vowels (e.g. in *peck*) with the phonological representation of the semantic prime vowel (e.g. *pack*). Based on the previous comparison between NZE and AE listeners in Author and Author (2010), the phonological representation of the TRAP vowel seems to be dialect-dependent and can account for the observed priming pattern.

Importantly, this study adds further evidence that the phonological, but not the phonetic representation determines the acceptance of vowel variants. This can be concluded from both the accuracy as well as the reaction time data.

Accuracies were significantly lower in the DRESS vowel condition and showed a trend to be lower in the KIT vowel condition than in the control condition. This suggests that lexical decisions were somewhat more difficult for the two conditions in which the stem vowel minimally differed from the stem vowel of the semantically related prime (with the TRAP vowel). We hypothesize that participants were uncertain whether or not to accept the KIT and DRESS vowels as variants of the TRAP vowel. Perhaps, this uncertainty was transferred onto the corresponding lexical decisions.

Reaction times showed that only semantically related primes facilitated the lexical decision latencies for their respective targets. Primes with the DRESS and KIT vowels (e.g. *peck*, *pick*) did not elicit significant priming (Figure 5, right). In fact, the pattern paralleled the outcome of the Author and Author (2010) study for NZE stimuli and AE listeners (Figure 5, middle). Note that the phonetic realization of the DRESS vowel in NZE corresponds to the realization of the KIT vowel in AE, while the realization of the KIT vowel in NZE is close to the realization of the DRESS vowel in AE (although more centralized in the vowel space).

Considering all three experiments - i.e. the two experiments of Author and Author (2010) and the experiment reported here - exemplar models cannot fully explain the patterns of results. While they can account for differences in vowel representations as elucidated before, and perhaps attribute the significant priming in the NZE pairs *peck-load* and *pick-load* to the inclusion of DRESS and KIT vowel variants in the exemplar set of the TRAP vowel, they would also expect target response latencies to depend on word frequencies or on the acoustic distance between variant vowels and the TRAP vowel. Neither of these correlations proved to be significant in the crucial test conditions of the three experiments. Furthermore, while fully specified models would be able to account for the pattern of this experiment, they fail to account for the significant priming if the prime vowel deviates from the TRAP vowel in NZE. We therefore argue that the overall pattern is only explicable by referring to differences in underlying phonological representations of the TRAP vowel between NZE and AE listeners on the basis of underspecified vs. specified vowel height. Accepted vowel variants are those whose extracted features do not mismatch with the abstract phonological representation of the vowel in the semantically related target word (here: TRAP words). We assume that non-mismatching variants of TRAP words co-activate the respective TRAP word and in turn lead to priming. More precisely, in NZE, accepted variants of the TRAP word were all those whose vowels did not mismatch with the underlying representation of the TRAP vowel. Since we assume that this vowel has no specification for tongue height, neither high nor mid vowels (from KIT or DRESS words) produced mismatches and equally led to co-activation of the TRAP words. Contrary, in AE, the TRAP vowel is specified for tongue height, i.e. is [low]. As a result, all (phonetically) non-low vowels (i.e. KIT and DRESS vowels) mismatched with the TRAP vowel, and correspondingly, the TRAP word variants could not prime its semantic relative. We therefore conjecture that long-term memory representations of vowel categories (described in more detail in Lahiri and Reetz, 2002 and Author and Author, 2010) are rather abstract, and sparsely specified for their contrasting features.

We acknowledge, however, that the priming pattern observed for the NZE listeners could be explained by alternative accounts that do not necessarily rely on abstract representations. For instance, it is conceivable that NZE listeners differ in their degree to rely on vowel identity from AE listeners in a lexical decision task, because the phonetic realizations of the NZE vowels show large variability compared to other English varieties¹. To exclude potential task effects that have nothing to do with underlying phonological representations, future research needs to provide means by which the current pattern can be replicated. In this regard, an experiment with AE materials given to NZE listeners might be potentially informative, although one has to bear in mind that NZE listeners would be quite familiar with AE accents from radio and TV. For the time being, we would like to argue that most parsimonious approach to optimally account for the observed pattern across the three experiments is given by the theory of underspecified, abstract long-term representations for speech sounds, as illustrated above.

4.2. Co-existence of abstract representations and phonetic details

We do not deny the fact that phonetic details play a role during the mapping of an acoustic speech signal onto its long-term memory representation. Previous research has shown that phonetic categories have internal structure with access to very detailed acoustic properties (cf. Allen and Miller, 2001; Miller, 1995). Additionally, a recent study provided evidence that while AE mid-vowels can be considered unspecified for height, the acoustic processing of these vowels still uses information relevant to vowel height (Author et al., 2012). The authors used the short front vowels [ɪ] and [æ] in a passive-oddball paradigm while the brain

¹This important aspect was brought to our attention by an anonymous reviewer.

electric response (Electroencephalogram, EEG) was recorded. In this paradigm, a repetitious presentation of one (standard) vowel was followed by a single presentation of a rare (deviant) vowel. This design elicits a typical even-related component in the EEG, the so-called Mismatch Negativity (MMN) as automatic auditory detection response to change or violation of rule-based predictions (Näätänen and Alho, 1997; Winkler, 2007). Importantly, this design has previously been employed to study different kinds of feature-based oppositions in vowels. Eulitz and Lahiri (2004) showed that an underspecified deviant vowel elicited an earlier and stronger MMN response if preceded by a more specific standard vowel than vice versa, i.e. if a more specific deviant vowel was preceded by an underspecified standard vowel. The assumption of the authors was that a more specific standard activates a more specific long-term memory representation for which the deviant represents a severe featural change. This assumption held for the comparison of underspecified coronal [ø] with specified dorsal [o], suggesting that acoustic information of coronality was processed in generating the MMN, while phonological information of coronality was unspecified in long-term memory. In the same vein, Author et al. (2010) expected to replicate these findings for tongue height oppositions [ɪ] vs. [æ] in American English. The low vowel [æ] in standard position was assumed to activate a more specific representation regarding tongue height, and the mid vowel [ɪ] should provide a mismatching acoustic height information. Vice versa, mid vowel [ɪ] in standard position was thought to activate an underspecified representation for which the deviant with low [æ] was more tolerable. Accordingly, the MMN in the former case ought to be of greater magnitude than in the latter case. This prediction was borne out by the results of Author et al. (2010), suggesting that even though mid vowels can be considered less specific in their long-term representation, acoustic information regarding their tongue height is not completely neglected in online-processing. This seems to be particularly true for a dense vowel system as found in AE. In its simplest definition, a dense vowel system is characterized by a high number of vowel categories within a given acoustic space (typically, the F1–F2 space), compared to a sparse system with only few vowel categories within the same space. Note that with respect to the experiment of Author et al. (2010), the experimentally observed MMN must be composite, combining abstract phonological category information with some detailed phonetic information.

To that end, the results of this study are partially compatible with exemplar models that stress the importance of fine-grained information in speech sound and word memory representations. Importantly, the co-existence of abstract and detailed information in current models of speech perception and representation is certainly not a neglected assumption (cf. Goldinger, 2007; Luce and Lyons, 1998; Luce et al., 2003). However, models differ as to whether they incorporate abstract representation in the model's architecture (Author and Author, 2010), or whether they regard abstract information as an emergent property (Ettlinger and Johnson, 2010; Hintzman, 1986; Luce and Lyons, 1998).

4.3. Exemplar models reconsidered

Previous support for exemplar-like or episodic representations predominantly stems from behavioral studies (for a review, see Tenpenny, 1995). Some of these studies focus on participants' sensitivity of within-category differences (e.g. McMurray et al., 2002; McMurray and Aslin, 2005). Other studies found that listeners retain non-linguistic information such as speaker gender, speaking rate and dialect in long-term memory (Church and Schacter, 1994; Pisoni, 1993; Schacter and Church, 1992). Mullenix et al. (1989) investigated the identification of English words, comparing a multi-talker condition with a single-talker condition. They obtained faster and more accurate responses in the single-talker than in the multiple-talker condition and interpreted these findings as evidence for the processing of non-linguistic information during speech perception. Similarly, Goldinger et

al. (1991) and Martin et al. (1989) showed that the recall performance on spoken stimuli is better if these stimuli are uttered by a single speaker than if they are uttered by multiple speakers. The same talker-variability effect emerged in the study by Bradlow et al. (1999), where the recognition accuracy of a spoken word was worse if it was repeated in a different voice than if it was repeated by the same voice. The effect of talker variability on speech perception led to the proposal that rather than normalizing the incoming speech signal, listeners directly map the acoustic percept onto very detailed lexical representations which are indexed for many non-linguistic factors (Johnson, 1997, 2006). Lexical representations are considered remembered exemplars or episodes with full phonetic details, which are consistently updated through experience. Since memory capacities are limited, not every encountered speech percept can be retained in long term memory and eventually decays if its occurrence is too infrequent (Pierrehumbert, 2001).

4.4. Integrative view

While studies seeking support for an exemplar-based lexicon suggest that extra-linguistic knowledge is based on implicit memory, it is not a necessary consequence to assume that these extra-linguistic properties are part of the concrete sound-meaning pairings in the mental lexicon. Even though non-linguistic variation in speech sounds can influence lexical access, it is important to ask at what stage of lexical processing and under which circumstances these factors play a role. For instance, although Goldinger (1996) replicated the speaker effect in a speeded classification task, he found a higher accuracy for the same voice in subjects' responses only for shallow levels of processing. The same-voice facilitation was less effective if participants had to engage deeper processing levels including access to syntactic category information. McLennan and Luce (2005) analyzed the extra-linguistic factors "speech rate" and "voice" in several long-term repetition priming experiments with lexical decision. Their results suggest that indexical variability (slow vs. fast speech, same vs. different voice) influences spoken word perception only when processing is slow and effortful. Differences between the processing of linguistic and extra-linguistic information were also found by Schweinberger and colleagues (Schweinberger, 2001; Schweinberger et al., 1997). Their findings suggest that speaker identification takes place at a different processing stage than phoneme recognition during speech recognition.

In sum, facilitating effects through extra-linguistic properties in priming experiments is not necessarily direct evidence for exemplars, but rather suggests that extra-linguistic properties might exert certain influences on speech perception at particular levels of processing. In this regard, neurophysiological and imaging research provides important counter-evidence for a very strong view of exemplar theory where linguistic (contrastive) and extra-linguistic (speaker-related) information is tightly interwoven. For instance, the research of Stevens (2004) showed that the neural systems for the memory of voices are different from those subserving word-recognition. Using fMRI, effects related to the processing of words involved left frontal and bilateral parietal cortical areas, whereas the processing of speaker voice was mainly visible in left temporal, right frontal and right medial parietal areas. Similar support for dissociable neural systems for words and voices, and in general, for dissociable subsystems for abstract linguistic versus specific extra-linguistic information has been provided by several neuroimaging studies (e.g. Formisano et al., 2008; Kreiman and van Lancker, 1988; Marsolek, 1999). Recently, Gonzalez and McLennan (2007) found that talker identity mismatches in spoken word recognition (specificity effects) affect the cerebral hemispheres differently. In particular, their long-term repetition experiments showed that there were specificity effects when stimuli were presented to the left ear but not when they were presented to the right ear. This means that the speaker identity had an influence for the right cerebral hemisphere, but not for the left hemisphere, supporting the view that extra-linguistic information is processed predominantly in the right hemisphere.

Thus, the overall evidence for a particular kind of exemplars where linguistic and extra-linguistic properties are tightly interwoven into very detailed and specific episodic representations is rather weak. Linguistic core representations in the brain are spatially dissociable from extra-linguistic properties, while the latter properties may influence and affect linguistic processing early on. Several ERP studies have shown that speaker voice attributes modulate early ERP components, suggesting that these attributes play an important role at early stages of speech perception (Beauchemin et al., 2006; Charest et al., 2009; Kreiman et al., 2005; Schweinberger, 2001; Titova and Näätänen, 2001). We therefore suggest that pure linguistic representations can be motivated by the fact that they are separable from non-linguistic representations at specific levels of processing, and at the level of cortical localization. We propose abstract long-term memory representations for vowels and presented evidence for such representations if targeting processing levels above pure acoustic discrimination. Exemplar models may better characterize processes at lower levels of processing where fine-grained acoustic information is vital. The challenge for future research is to link or to extend abstract models in a way that they can account for both specificity and categorical effects (e.g. hybrid models). We believe that the FUL model is a promising candidate in this respect, because it embraces both levels of processing in its architecture and its assumption regarding the mapping of acoustic information onto long-term (abstract) memory representations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Agresti, A. Categorical data analysis. Wiley; New York: 2002.
- Allen JS, Miller JL. Contextual influences on the internal structure of phonetic categories: A distinction between lexical status and speaking rate. *Perception & Psychophysics*. 2001; 63:798–810. [PubMed: 11521848]
- Archangeli D. Aspects of underspecification theory. *Phonology*. 1988; 5:183–207.
- Author, 2009.
- Author, Author, 2010.
- Author, et al.
- Baayen H, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory & Language*. 2008; 59:390–412.
- Baayen, H.; Piepenbrock, R.; Gulikers, L. Linguistic Data Consortium. University of Pennsylvania; Philadelphia, PA: 1995. The CELEX Lexical Database [CD-Rom].
- Bauer L. Notes on New Zealand English phonetics and phonology. *English World-Wide: A Journal of Varieties of English*. 1986; 7:225–258.
- Bauer, L. Spelling pronunciation and related matters in New Zealand English. In: Lewis, JW., editor. *Studies in General and English Phonetics: Essays in Honour of Professor J. D. O'Connor*. Routledge; London: 1995. p. 320-325.
- Beauchemin M, Beaumont Ld, Vannasing P, Turcotte A, Arcand C, Belin P, Lassonde M. Electrophysical markers of voice familiarity. *European Journal of Neuroscience*. 2006; 23:3081–3086. [PubMed: 16819998]
- Bell A. The phonetics of fish and chips in New Zealand: Marking national and ethnic identities. *English World-Wide: A Journal of Varieties of English*. 1997; 18:243–270.
- Benguerel AP, McFadden TU. The effect of coarticulation on the role of transitions in vowel perception. *Phonetica*. 1989; 46:80–96. [PubMed: 2608727]
- Blumstein SE, Stevens KN. Phonetic features and acoustic invariance in speech. *Cognition*. 1981; 10:25–32. [PubMed: 7198546]

- Boersma, P.; Weenink, D. PRAAT: Doing Phonetics by Computer (ver. 5.2.24). Institut for Phonetic Sciences; Amsterdam: 2011.
- Bradlow AR, Nygaard LC, Pisoni DB. Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics*. 1999; 61:206–219. [PubMed: 10089756]
- Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*. 1993; 88:9–25.
- Browman CP, Goldstein L. Articulatory Phonology: An overview. *Phonetica*. 1992; 49:155–180. [PubMed: 1488456]
- Bybee, J. *Phonology and Language Use*. Cambridge University Press; Cambridge, UK: 2001.
- Bybee, J.; Hopper, PJ., editors. *Frequency and the emergence of linguistic structure*. John Benjamins; Amsterdam: 2001.
- Calabrese A. A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry*. 1995; 26:373–463.
- Charest I, Pernet CR, Rousselet G, Quinones I, Latinus M, Fillion-Bilodeau S, Chartrand J-P, Belin P. Electrophysiological evidence for an early processing of human voices. *BMC Neuroscience*. 2009; 12710.1186/1471-2202-1110-1127
- Chomsky, N.; Halle, M. *The Sound Pattern of English*. Harper and Row; New York: 1968.
- Church BA, Schacter DL. Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 1994; 20:521–533.
- Crawley, MJ. *Statistics. An Introduction using R*. Wiley; Chichester: 2005.
- Cutler A. The comparative perspective on spoken-language processing. *Speech Communication*. 1997; 21:3–15.
- Daelemans, W.; Van den Bosch, A. *Memory-Based Language Processing*. Cambridge University Press; 2005.
- Diehl RL. Searching for an auditory description of vowel categories. *Phonetica*. 2000; 57:267–274. [PubMed: 10992146]
- Diehl, RL.; Kluender, KR.; Walsh, MA.; Parker, EM. Auditory enhancement in speech perception and phonology. In: Hoffman, R.; Palermo, DS., editors. *Cognition and the Symbolic Processes: Applied and Ecological Perspectives*. Lawrence Erlbaum; Hillsdale, NJ: 1991.
- Dudani SA. The distance-weighted k-Nearest-Neighbor rule. *IEEE Transactions on Systems, Man and Cybernetics*. 1976; SMC-6:325–327.
- Easton A, Bauer L. An acoustic study of the vowels of New Zealand English. *Australian Journal of Linguistics: Journal of the Australian Linguistic Society*. 2000; 20:93–117.
- Ernestus, M. Unpublished Dissertation. University of Amsterdam; Amsterdam: 2000. Voice assimilation and segment reduction in casual Dutch.
- Ernestus M, Lahey M, Verhees F, Baayen RH. Lexical frequency and voice assimilation. *Journal of the Acoustical Society of America*. 2006; 120:1040–1051. [PubMed: 16938990]
- Ettlinger M, Johnson K. Vowel Discrimination by English, French and Turkish speakers: Evidence for an exemplar-based approach to speech perception. *Phonetica*. 2010; 66:222–242. [PubMed: 20431322]
- Eulitz C, Lahiri A. Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive Neuroscience*. 2004; 16:577–583. [PubMed: 15185677]
- Fant, G. *Acoustic Theory of Speech Production*. Mouton de Gruyter; The Hague: 1960.
- Flege JE, Munro MJ, Fox RA. Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America*. 1994; 95:3623–3641. [PubMed: 8046152]
- Flemming E. Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology*. 2001; 18:7–44.
- Formisano E, De Martino F, Bonte ML, Goebel R. “Who” is saying “what”? Brain-based decoding of human voice and speech. *Science*. 2008; 322:970–973. [PubMed: 18988858]
- Forster KI. The microgenesis of priming effects in lexical access. *Brain and Language*. 1999; 68:5–15. [PubMed: 10433733]

- Forster KI, Forster JC. DMDX: a windows display program with millisecond accuracy. *Behavior Research Methods, Instruments and Computers*. 2003; 35:116–124.
- Fowler CA. Parsing coarticulated speech in perception: Effects of coarticulation resistance. *Journal of Phonetics*. 2005; 33:199–213.
- Gaskell GM, Marslen-Wilson WD. Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*. 1996; 22:144–158. [PubMed: 8742258]
- Goldinger SD. Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1996; 22:1166–1183.
- Goldinger SD. Echoes of echoes? An episodic theory of lexical access. *Psychological Review*. 1998; 105:251–279. [PubMed: 9577239]
- Goldinger, SD. A complementary-system approach to abstract and episodic speech perception. Paper presented at the Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS); Saarbrücken. 2007. p. 49-54.
- Goldinger SD, Pisoni DB, Logan JS. On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 1991; 17:152–162.
- Gonzalez J, McLennan CT. Hemispheric differences in indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 2007; 33:410–424.
- Gordon, E.; Maclagan, M.; Hay, J.; Campbell, L.; Trudgill, P. *New Zealand English: Its Origins and Evolution*. Cambridge University Press; Cambridge, UK: 2004.
- Gove, PB. *Webster's New Dictionary of Synonyms*. Merriam; Springfield, MA: 1968.
- Gow DW Jr. Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*. 2001; 45:133–159.
- Hagiwara R. Dialect variation and formant frequency: The American English vowels revisited. *The Journal of the Acoustical Society of America*. 1997; 102:655–658.
- Halle M. On the origin of the distinctive features. *International Journal of Slavic Linguistics and Poetics*. 1983; 27:77–86.
- Halle M, Vaux B, Wolfe A. On feature spreading and the representation of place of articulation. *Linguistic Inquiry*. 2000; 31:387–444.
- Hasselmo, ME. *Brain Mechanisms of Episodic Memory*. MIT Press; Cambridge, MA: 2012. *How We Remember*.
- Hawkins P. A phonemic transcription system for New Zealand English. *Te Reo: Journal of the Linguistic Society of New Zealand*. 1973; 16:15–21.
- Hawkins P. The role of NZ English in a binary feature analysis of English short vowels. *Journal of the International Phonetic Association*. 1976; 6:50–66.
- Hay J, Nolan A, Drager K. From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review*. 2006; 23:351–379.
- Hillenbrand J, Getty LA, Clark MJ, Wheeler K. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*. 1995; 97:3099–3111. [PubMed: 7759650]
- Hintzman DL. "Schema Abstraction" in a Multiple-Trace Memory Model. *Psychological Review*. 1986; 93:411–428.
- Inkelas, S. The consequences of optimization for underspecification. In: Buckley, E.; Iatridou, S., editors. *Proceedings of the Twenty-Fifth Northeastern Linguistics Society*. GLSA; Amherst, MA: 1995. p. 287-302.
- Itô J, Mester A, Padgett J. Licensing and underspecification in optimality theory. *Linguistic Inquiry*. 1995; 26:571–613.
- Jaeger FT. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory & Language*. 2008; 59:434–446. [PubMed: 19884961]
- Jakobson, R. Observations sur le classement phonologique des consonnes. *Proceedings of the 3rd International Conference of Phonetic Sciences*; Ghent. 1939. p. 34-41.

- Jakobson, R.; Fant, G.; Halle, M. The distinctive features and their correlates. The MIT Press; Cambridge, MA: 1952. Preliminaries to speech analysis.
- Johnson, K. Speech perception without speaker normalization. In: Johnson, K.; Mullennix, JW., editors. Talker Variability in Speech Processing. Academic Press; New York: 1997. p. 145-166.
- Johnson, K. Speaker normalization in speech perception. In: Pisoni, DB.; Remez, RE., editors. The handbook of speech perception. Blackwell Publishing; Malden (MA): 2005. p. 363-389. et al
- Johnson K. Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*. 2006; 34:485–499.
- Jun, J. Place assimilation. In: Hayes, B.; Kirchner, R.; Steriade, D., editors. Phonetically based phonology. Cambridge University Press; Cambridge: 2004. p. 58-86.
- Kewley-Port D. Time-varying features as correlates of place of articulation in stop consonants. *Journal of the Acoustical Society of America*. 1983; 73:322–335. [PubMed: 6826902]
- Kreiman J, van Lancker D. Hemispheric specialization for voice recognition: Evidence from dichotic listening. *Brain and Language*. 1988; 34:246–252. [PubMed: 3401693]
- Kreiman, J.; Vanlancker-Sidtis, D.; Gerratt, BR. Perception of voice quality. In: Pisoni, DB.; Remez, RE., editors. The Handbook of speech perception. Blackwell Publishing; Malden (MA): 2005. p. 338-362. et al
- Labov, W.; Ash, S.; Boberg, C. Atlas of North American English. Mouton de Gruyter; Berlin: 2006.
- Ladefoged, P. Vowels and Consonants: An Introduction to the Sounds of Languages. Blackwell; Malden, MA: 2001.
- Lahiri A, Gewirth L, Blumstein SE. A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: evidence from a cross-language study. *Journal of the Acoustical Society of America*. 1984; 76:391–404. [PubMed: 6480990]
- Lahiri, A.; Reetz, H. Underspecified recognition. In: Gussenhoven, C.; Warner, N., editors. Laboratory Phonology VII. Mouton de Gruyter; Berlin: 2002. p. 637-677.
- Lahiri A, Reetz H. Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics*. 2010; 38:44–59.
- Langstrof C. The short front vowels in New Zealand English in the Intermediate Period. *New Zealand English Journal*. 2003; 17:4–16.
- Langstrof C. Acoustic evidence for a push-chain shift in the Intermediate Period of New Zealand English. *Language Variation and Change*. 2006; 18:141–164.
- Lindblom, B. Explaining phonetic variation: A sketch of the H&H theory. In: Hardcastle, WJ.; Marchal, A., editors. Speech Production and Speech Modelling. Kluwer Academic Publishers; Dordrecht: 1990. p. 403-439.
- Lindblom BEF, Studdert-Kennedy M. On the role of formant transitions in vowel recognition. *The Journal of the Acoustical Society of America*. 1967; 42:830–830. [PubMed: 6075568]
- Lively, SE.; Pisoni, DB.; Goldinger, SD. Spoken word recognition. In: Gernsbacher, MA., editor. Handbook of Psycholinguistics. Academic Press; San Diego: 1994. p. 265-301. et al
- Lombardi L. Postlexical rules and the status of privative features. *Phonology*. 1996; 13:1–38.
- Luce PA, Lyons EA. Specificity of memory representations for spoken words. *Memory & Cognition*. 1998; 26:708–715. [PubMed: 9701963]
- Luce, PA.; McLennan, CT.; Charles-Luce, J. Abstractness and specificity in spoken word recognition: Indexical and allophonic variability in long-term repetition priming. In: Bowers, JS.; Marsolek, CJ., editors. Rethinking implicit memory. Oxford University Press; Oxford, New York: 2003. p. 197-214.
- Maclagan MA, Gordon E. The story of New Zealand English: What the ONZE project tells us. *Australian Journal of Linguistics: Journal of the Australian Linguistic Society*. 2004; 24:41–56.
- Maclagan, MA.; Hay, J. The rise and rise of New Zealand English DRESS. Paper presented at the 10th Australian International Conference on Speech Science and Technology; Sydney: Macquarie University; 2004. p. 183-188.
- Marslen-Wilson WD, Nix A, Gaskell GM. Phonological variation in lexical access: Abstractness, inference and English place assimilation. *Language and Cognitive Processes*. 1995; 10:285–308.

- Marsolek CJ. Dissociable neural subsystems underlie abstract and specific object recognition. *Psychological Science*. 1999; 10:111–118.
- Martin CS, Mullenix JW, Pisoni DB, Summers WV. Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 1989; 15:676–684.
- McCarthy JJ, Taub A. Review of Paradis and Prunet (1991). *Phonology*. 1992; 9:363–370.
- McLennan CT, Luce PA. Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 2005; 31:306–321.
- McMurray B, Aslin RN. Infants are sensitive to within-category variation in speech perception. *Cognition*. 2005; 95:B15–B26. [PubMed: 15694642]
- McMurray B, Tanenhaus MK, Aslin RN. Gradient effects of within-category phonetic variation on lexical access. *Cognition*. 2002; 86:B33–B42. [PubMed: 12435537]
- Meyer D, Schvaneveldt RW. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*. 1971; 90:227–234. [PubMed: 5134329]
- Miller JL. On the internal structure of phonetic categories: A progress report. *Cognition*. 1995:333–347.
- Mitterer H, Ernestus M. Listeners recover /t/s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics*. 2006; 34:73–103.
- Mitterer H, Yoneyama K, Ernestus M. How we hear what is hardly there: Mechanisms underlying compensation for /t/-reduction in speech comprehension. *Journal of Memory and Language*. 2008; 59:133–152.
- Monahan P, Idsardi WJ. Auditory sensitivity to formant ratios: Toward an account of vowel normalization. *Language and Cognitive Process*. 2010; 25:808–839.
- Mullenix JW, Pisoni DB, Martin CS. Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*. 1989:85.
- Näätänen R, Alho K. Mismatch negativity (MMN) - the measure for central sound representation accuracy. *Audiology and Neuro-Otology*. 1997; 2:341–353. [PubMed: 9390839]
- Nosofsky RM, Palmeri TJ. An exemplar-based random walk model of speeded classification. *Psychological Review*. 1997; 104:266–300. [PubMed: 9127583]
- Nosofsky RM, Zaki SR. Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2002; 28:924–940.
- Paradis, C.; Prunet, J-F. *The Special Status of Coronals: Internal and External Evidence*. Academic Press; San Diego: 1991.
- Peterson GE, Barney HL. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*. 1952; 24:175–184.
- Pierrehumbert, J. Word-specific phonetics. In: Gussenhoven, C.; Warner, N., editors. *Laboratory Phonology VII*. Mouton de Gruyter; Berlin: 2002. p. 101-140.
- Pierrehumbert, JB. Exemplar Dynamics: Word frequency, lenition and contrast. In: Bybee, J.; Hopper, PJ., editors. *Frequency and the emergence of linguistic structure*. John Benjamins; Amsterdam: 2001. p. 137-158.
- Pinheiro, JC.; Bates, DM. *Mixed-effects models in S and S-PLUS*. Springer Verlag; 2000.
- Pisoni DB. Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics*. 1973; 13:253–260. [PubMed: 23226880]
- Pisoni DB. Long-term memory in speech perception: Some new findings on talker variability, speaking rate and perceptual learning. *Speech Communication*. 1993; 13:109–125. [PubMed: 21461185]
- Pisoni, DB. Some thoughts on “normalization” in speech perception. In: Johnson, K.; Mullenix, JW., editors. *Talker Variability in Speech Processing*. Academic Press; San Diego: 1997. p. 9-32.
- Pols LCW, van der Kamp LJT, Plomp R. Perceptual and physical space of vowel sounds. *The Journal of the Acoustical Society of America*. 1969; 46:458–458. [PubMed: 5804118]

- Prince, A.; Smolensky, P. *Optimality Theory: Constraint interaction in Generative Grammar*. Blackwell; Malden, MA: 2004.
- Raymond WD, Dautricourt R, Hume E. Word internal /t,d/ deletion in spontaneous speech: Modelling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change*. 2006; 18:55–97.
- Schacter DL, Church BA. Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 1992; 18:915–930.
- Schouten MEH, van Hessen AJ. Modeling phoneme perception: Categorical perception. *The Journal of the Acoustical Society of America*. 1992; 92:1841–1855. [PubMed: 1401529]
- Schweinberger SR. Human brain potential correlates of voice priming and voice recognition. *Neuropsychologia*. 2001; 39:921–936. [PubMed: 11516445]
- Schweinberger SR, Herholz A, Sommer W. Recognizing famous voices: The influence of stimulus duration and different types of retrieval cues. *Journal of Speech, Language, and Hearing Research*. 1997; 40:453–463.
- Snoeren ND, Seguí J, Hallé PA. On the role of regular phonological variation in lexical access: Evidence from voice assimilation in French. *Cognition*. 2008; 108:512–521. [PubMed: 18400217]
- Stevens AA. Dissociating the cortical basis of memory for voices, words and tones. *Brain Research. Cognitive Brain Research*. 2004; 18:162–171. [PubMed: 14736575]
- Stevens, K. *Acoustic Phonetics*. The MIT Press; Cambridge, MA: 1998.
- Stevens K, Blumstein SE. Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*. 1978; 64:1358–1368. [PubMed: 744836]
- Stevens KN. Toward a model for lexical access based on acoustic landmarks and distinctive features. *The Journal of the Acoustical Society of America*. 2002; 111:1872–1891. [PubMed: 12002871]
- Sussman HA, McCaffrey HA, Matthews SA. An investigation of locus equations as a source of relational invariance for stop place categorization. *Journal of the Acoustical Society of America*. 1991; 90:1309–1325.
- Tenpenny PL. Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin and Review*. 1995; 2:339–363. [PubMed: 24203715]
- Thomas, B. Unpublished Master Thesis. University of Canterbury; Christchurch: 2004. In Support of an Exemplar-Based Approach to Speech Perception and Production: A Case Study on the Merging of Pre-Lateral DRESS and TRAP in New Zealand English.
- Thomas ER. An acoustic analysis of vowel variation in New World English. *American Speech*. 2001; 85:1–14.
- Titova N, Näätänen R. Preattentive voice discrimination by the human brain as indexed by the mismatch negativity. *Neuroscience Letters*. 2001; 308:63–65. [PubMed: 11445287]
- Warren, P.; Hay, J. Using sound change to explore the mental lexicon. In: Haberman, G.; Fletcher-Flinn, C., editors. *Cognition and Language: Perspectives from New Zealand*. Australian Academic Press; Brisbane: 2006. p. 105-126.
- Warren, P.; Hay, J.; Thomas, B. The loci of sound change effects in recognition and perception. In: Cole, J.; Hualde, JI., editors. *Laboratory Phonology IX*. Mouton de Gruyter; New York: 2008. p. 88-112.
- Watson CI, Harrington J, Evans Z. An acoustic comparison between New Zealand and Australian English vowels. *Australian Journal of Linguistics: Journal of the Australian Linguistic Society*. 1998; 18:185–207.
- Watson CI, Maclagan M, Harrington J. Acoustic evidence for vowel change in New Zealand English. *Language Variation and Change*. 2000; 12:51–68.
- Wells, JC. *Accents of English I: Introduction*. Cambridge University Press; Cambridge, UK: 1982a.
- Wells, JC. *Accents of English III: Beyond the British Isles*. Cambridge University Press; Cambridge, UK: 1982b.
- Winkler I. Interpreting the Mismatch Negativity. *Journal of Psychophysiology*. 2007; 21:147–163.
- Zimmerer, F.; Reetz, H.; Lahiri, A. Harmful reduction?. Paper presented at the Laboratory Phonology; Wellington (NZ). 2008.

Highlights

- American English is restrictive in height variation of front vowels
- This contrasts with New Zealand English with a front vowel shift
- Abstract vowel representations can account for both findings
- Sparse and detailed representations depend on the level of processing

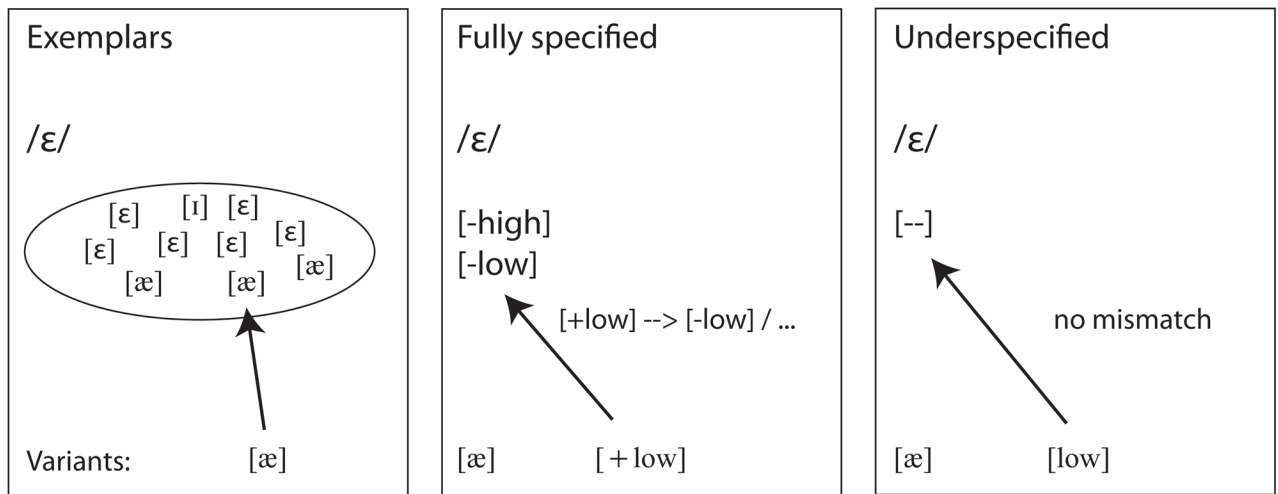


Figure 1.

Comparison of different representational models regarding the toleration of a mid vowel variant that crosses a categorical boundary. Left: Exemplar models include vowel variants directly in the exemplar sets, and the mapping can occur directly. Middle: Fully specified featural models have to stipulate a licensing rule that would allow a low vowel to be interpreted as a mid vowel. Right: The underspecified featural model can directly map a low vowel onto the underspecified representation of a mid vowel.

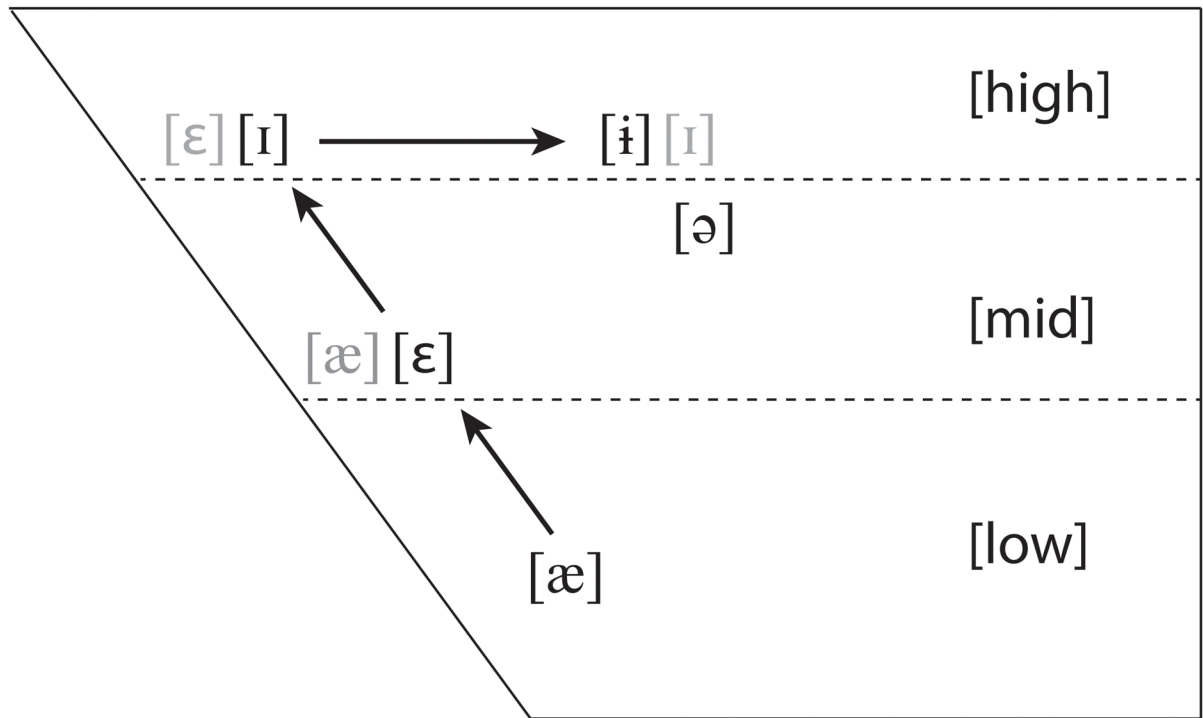


Figure 2. NZE short front vowel shift. Locations of the AE vowels are given in black, while the endpoints of the shift in NZE are shown in grey.

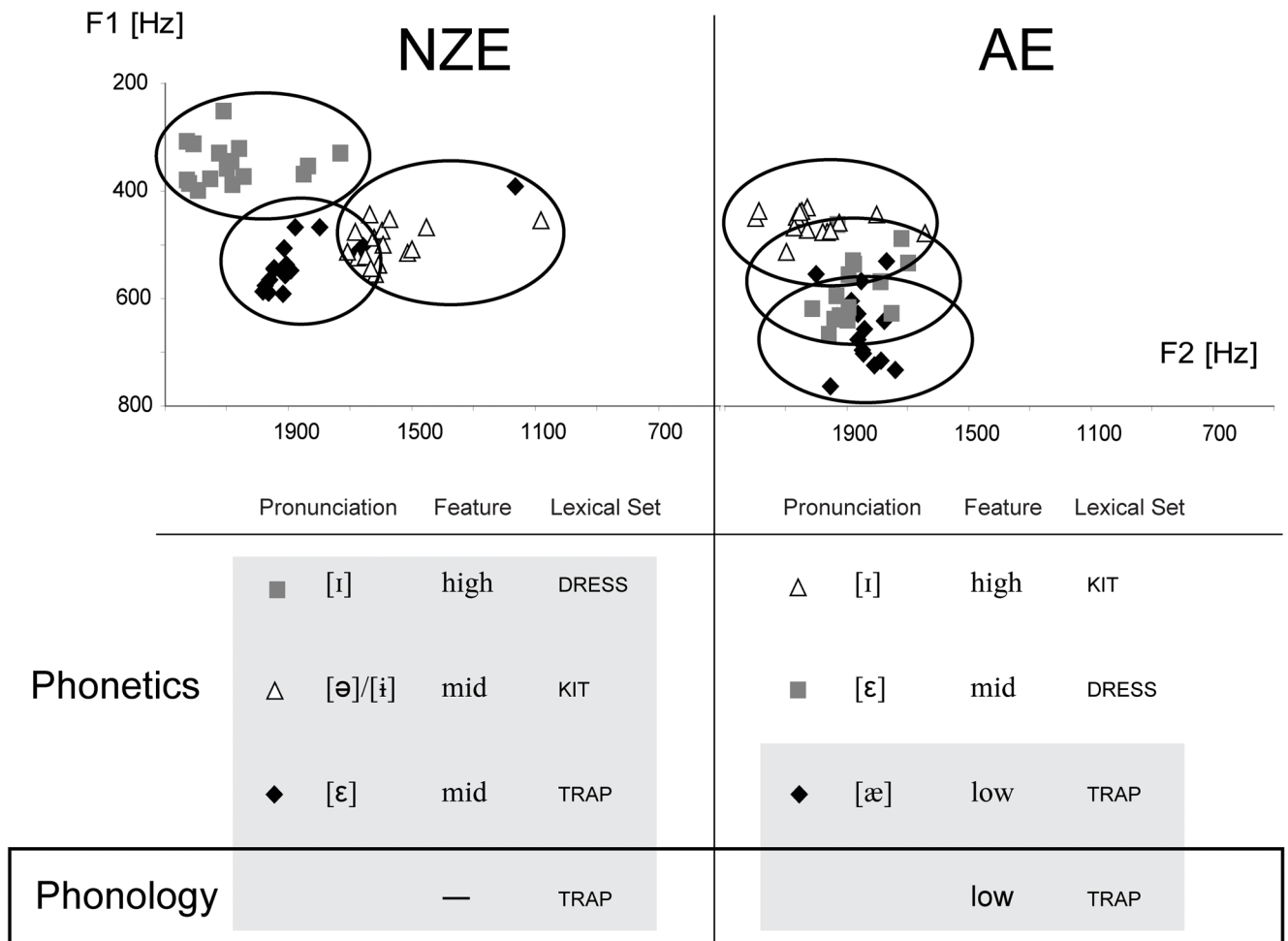


Figure 3. Phonetic realization of experimental stimuli in Author and Author (2010) [NZE] and this study [AE]. The top part shows the location of the vowels from monomorphemic English nouns in the F1/F2 space. The bottom part illustrates the assumed phonetic and phonological categorizations. Lexical set labels are given in order to compare the categories across dialects. Accepted variants (shaded in grey) are determined on the basis of feature agreements (NZE) that are absent when the phonological representation has a specification for tongue height (AE).

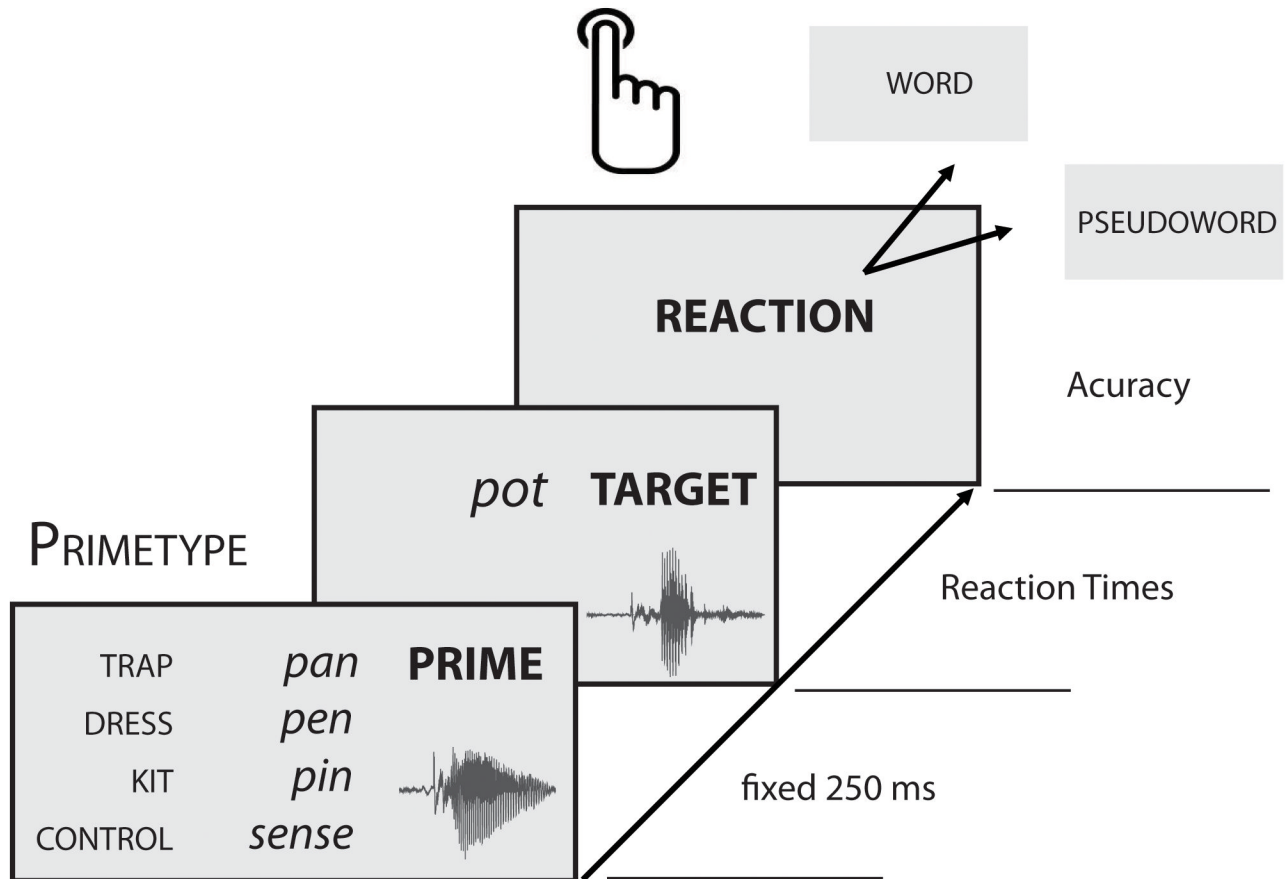


Figure 4. Experimental design of auditory semantic priming. Primetype refers to the lexical set by which the vowel in the prime words is described. The control condition comprises a prime without any relation to the target.

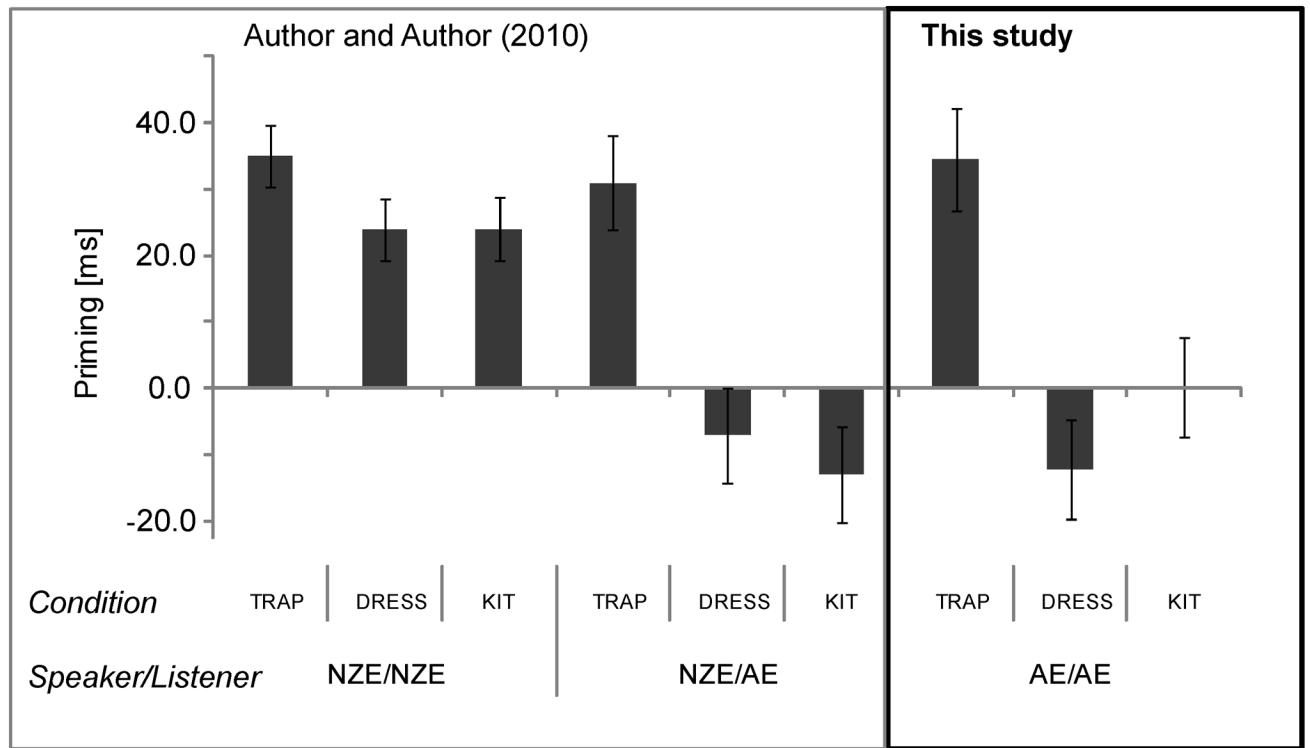


Figure 5. Amount of priming in the three vowel conditions (difference between reaction times in control and vowel conditions), compared to the Author and Author (2010) findings. Error bars indicate standard errors of the mean.

Table 1

Mixed-Effect Logit Model on accuracy, using treatment coding for the fixed effect condition with the control condition (unrelated) as reference. The dependent variable is dichotomous with $y=1$ (correct) and $y=0$ (incorrect).

Fixed	Estimate	Std. Error	z value	$p(> z)$
(Intercept)	3.9176	0.4457	8.789	0.0000
Prime Type: TRAP	0.4104	0.4906	0.837	0.4028
Prime Type: DRESS	-0.9117	0.3956	-2.304	0.0212
Prime Type: KIT	-0.7434	0.4001	-1.858	0.0632

Table 2

Fixed effects for the Mixed Model on reaction times. Fixed effects are given in relation to the control condition. The p values were calculated using 10000 Markov Chain Monte Carlo simulations (MCMC; Baayen, 2008).

Fixed	Estimate	Std. Error	MCMCmean	pMCMC	t value	p(> t)
(Intercept)	6.8524	0.0301	6.8518	0.0001	227.8900	0.0000
Prime Type: TRAP	-0.0335	0.0152	-0.0340	0.0338	-2.2000	0.0280
Prime Type: DRESS	0.0146	0.0153	0.0145	0.3616	0.9500	0.3404
Prime Type: KIT	0.0042	0.0153	0.0036	0.8152	0.2700	0.7848