Available online at www.sciencedirect.com

ScienceDirect

Lingua 140 (2014) 35–51

Lingua

www.elsevier.com/locate/lingua

Sparseness of vowel category structure: Evidence from English dialect comparison

Mathias Scharinger [a,b,*], William J. Idsardi [b]

[a] *Max Planck Research Group ''Auditory Cognition'', Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*

[b] *Department of Linguistics, University of Maryland, USA*

Received 21 March 2012; received in revised form 19 November 2013; accepted 22 November 2013
Available online 4 January 2014

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 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

* Corresponding author at: Max Planck Research Group ''Auditory Cognition'', Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1A, 04103 Leipzig, Germany. Tel.: +49 0341 9940 2482.

E-mail address: mscharinger@cbs.mpg.de (M. Scharinger).

0024-3841/$ – see front matter
http://dx.doi.org/10.1016/j.lingua.2013.11.007

impressions are strengthened by the findings that vowel perception is less categorical than (stop) consonant perception (Pisoni, 1973; Schouten and van Hessen, 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 (Scharinger and Lahiri, 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 (Scharinger 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, 2001, 2002), 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 Fig. 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, 2001, 2002; Pisoni, 1997; Thomas, 2004). As a consequence of these architectural assumptions, there are two crucial measures that should

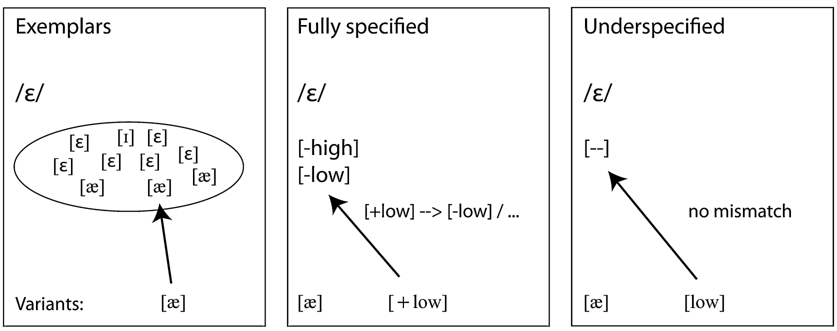

Fig. 1. Comparison of different representational models regarding the toleration of a mid vowel variant that crosses a categorical boundary. Left: exemplar models include vowel variants directly in the exemplar sets, and the mapping can occur directly. Middle: fully specified featural models have to 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.