Journal of Phonetics 45 (2014) 43–51

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Acoustic correlates of allophonic versus phonemic dimensions in monolingual and bilingual infants' input

Amanda Seidl [a], Kristine H. Onishi [b], Golnoush Alamian [b,c], Alejandrina Cristia [d,*]

[a] Purdue University, West Lafayette, IN 47907, USA
[b] McGill University, Montreal, QC, Canada H3A 1B1
[c] University of British Columbia, Vancouver, BC, Canada V6T 1Z4
[d] Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS, DEC-IEC-ENS-PSL*, EHESS Paris 75005, France

ARTICLE INFO

Article history:
Received 17 September 2013
Received in revised form
12 February 2014
Accepted 5 March 2014
Available online 12 April 2014

Keywords:
Allophones
Bilingualism
Infancy
Vowel nasalization
Vowel tenseness

ABSTRACT

Allophones are diverse phonetic instantiations of a single underlying sound category. As such, they pose a peculiar problem for infant language learners: These variants occur in the ambient language, but they are not used to encode lexical contrasts. Infants' sensitivity to sounds varying along allophonic dimensions declines by 11 months of age, suggesting that there must be information to phonological status available to pre-lexical infants. The present work tests one specific type of information: acoustic implementation. It was hypothesized that the acoustic distance between two vowel categories is smaller when the dimension along which the two vowels differ is allophonic (e.g., vowel nasality in American English, vowel tenseness in Quebec French) compared to when it is phonemic (e.g., vowel tenseness in American English, vowel nasality in Quebec French). Monolingual mothers speaking either English or French and bilingual mothers speaking both languages were recorded while they described objects to their 11-month-olds. Results provided weak support for the main hypothesis.

1. Introduction

By the end of the first year of life, both monolingual and bilingual infants have begun to lay down the foundations for their native language(s). One of the key landmarks of this process involves perceptual attunement, the process by which infants allocate their attention more effectively by reducing sensitivity to non-native sound distinctions and increasing it for native ones (see Tsuji & Cristia, 2014, for a review). Recent work suggests that the decline even applies to pairs of sounds that are present in the input, but which differ on a dimension that is allophonic in the native language. Indeed, English-learning 11-month-olds fail to learn a pattern involving vowel nasality, a dimension that is allophonic in the ambient language, whereas both 4-month-old English learners and French-learning 11-month-olds succeed (Seidl, Cristia, Bernard, & Onishi, 2009). How might they achieve this early feat? Here, we evaluate a heretofore ignored potential contributor: the acoustic separation of sound categories. Specifically, we measure acoustic separation as a function of phonemic or allophonic status in infants' input.

Although there is increasing understanding of more intermediate cases (Hall, 2013), the classical description of allophony involves two clear scenarios. A pair of sounds are allophones of the same phoneme if they are in complementary distribution (for example, one variant occurs in one phonological environment, the other variant in all other environments) or in free variation (either is acceptable in a given context, and no meaning change ensues from swapping them; Kenstowicz, 1994). Children could use at least three strategies to learn about phonological status. The first two have been studied in previous research, which we summarize briefly before turning to the one focused in the present work.

The first and obvious strategy involves *lexical bootstrapping*. Hearing [tɛlɪtʌbi] and [tɛlɪd̰ʌbi] spoken in reference to the same Teletubby toy could indicate that the voiceless [t] and glottalized [d̰] alveolar stops are in free variation (Foulkes, Docherty, & Watt, 2005). Recent modeling work suggests that a similar strategy of mapping word forms to meanings or referents improves the determination of allophonic status even when applied pre-lexically, using 'pseudowords' (i.e., consistent sequences of sounds that may or may not correspond to actual words of the language; Fourtassi & Dupoux, in preparation; Martin, Peperkamp, & Dupoux, 2013). According to these proposals, infants would not need to have large semantic vocabularies to profit from lexical-type cues to allophony. As of yet, there is no experimental work directly supporting infants' *use* of such cues.

Second, *distributional information* is extremely informative to determine allophones in complementary distribution. The possibility that complementary distribution would affect perception is strengthened by evidence of reduced sensitivity in 12-month-olds after experimental exposure

* Corresponding author. Tel.: +33 1 44 32 26 23; fax: +33 1 44 32 26 30.
E-mail addresses: aseidl@purdue.edu (A. Seidl), kris.onishi@mcgill.ca (K.H. Onishi), g.alamian@alumni.ubc.ca (G. Alamian), alecristia@gmail.com (A. Cristia).

0095-4470/$ - see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2014.03.004

to sounds in complementary distribution (White, Peperkamp, Kirk, & Morgan, 2008). However, modeling work demonstrates that distributional information alone is not sufficient to establish which sounds are allophones of the same phoneme when complementary distribution affects whole sound classes. To take the example of vowel nasalization in English, a learner with access to distributional information alone cannot decide whether [i] and [ĩ] map onto the same phoneme and [i] and [ũ] to different ones, or whether [i] and [ũ] are allophones instead. Peperkamp, LeCalvez, Nadal, and Dupoux (2006) document precisely this behavior in a computational model having access only to complementary distribution. They also demonstrate that performance is improved when the learner constrained hypothetical allophones on the basis of phonetic similarity. In the next subsection, we explain how acoustic information on its own could potentially facilitate learning of phonological status in pre-lexical infants.

1.1. Phonetic instantiation as an additional cue to phonological status

A great deal of research in psychology suggests that category structure affects similarity judgments (e.g., Rosch, 1975). This general principle could also be evident in the phonetics–phonology interface (e.g., Kingston, 2007). A pair of sounds could come to be perceptually, articulatorily, and acoustically less different from each other if the dimension along which the sounds contrast is allophonic compared to when it is phonemic through the following perception–production loop.

First, a host of evidence shows that phonemic distinctions are more salient than comparable allophonic distinctions for adult perceivers. (For brevity, we refer to distinctions between allophones of the same phoneme as *allophonic* distinctions and distinctions between sounds that map onto different phonemic categories as *phonemic* distinctions.) To begin with, adult listeners rate sounds as more similar to each other if the sounds are used allophonically in the listeners' native language than if they are used phonemically (e.g., Boomershine, Hall, Hume, & Johnson, 2008; Johnson & Babel, 2010). Moreover, discrimination is faster, more accurate, and more categorical when a distinction is phonemic than when it is allophonic (Beddor & Strange, 1982; Boomershine et al., 2008; Harnsberger, 2001; Whalen, Best, & Irwin, 1997). Finally, sounds in an allophonic distinction do not block word recognition (e.g., the clear pronunciation of 'atom' [ætəm] and the flapped pronunciation [ærəm] prime each other and themselves to a similar extent, whereas sounds in a phonemic distinction do not, e.g., 'Adam' [ædəm] does not prime [ætəm]; McLennan, Luce, & Charles-Luce, 2003).

These perceptual differences could lead to less clear productions of allophonic distinctions since speakers tend to articulate less clearly distinctions that they find harder to discriminate (Perkell et al., 2004). Furthermore, given that listeners can still retrieve the lexical item even when a different allophone is uttered (McLennan et al., 2003), there is less pressure for talkers to hyperarticulate allophonic distinctions (Lindblom, 1990).

At this point, articulation feeds back on perception: If speakers hypo-articulate, then allophonic distinctions will tend to be acoustically less distinct than comparable phonemic distinctions. Indeed, such patterns have sometimes been reported (e.g., Gick, Pulleyblank, Campbell, & Mutaka, 2006, report that, in Kinande, the high vowels, phonemically contrastive in tenseness or advanced tongue root [ATR], are articulatorily more distinct than the low vowels that differ only allophonically in ATR; see also Johnson & Babel, 2010; Spears, 2006; Ussishkin & Wedel, 2009). If sounds come to be physically closer in the acoustic input of these speakers, then they will become perceptually more similar, reinforcing the perceptual step noted above.

Finally, the effects of this feedback loop could come to be 'phonologized' historically, such that sounds that are in an allophonic distinction would tend to become more similar to each other acoustically with passing generations. Regardless of whether phonological status affects acoustic implementation only synchronically, or whether additionally its effects come to be phonologized, the acoustic separation between two sounds would be smaller when they are in an allophonic, compared to when they are in a phonemic, distinction, in the spoken input available to the infant learner.

It is reasonable to suppose that such fine differences in acoustic instantiation could in turn affect infants' perception, since their perceptual categories are altered by the distributions of acoustic cues they encounter (Cristia, 2011; Cristia, McGuire, Seidl, & Francis, 2011; Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002; Yoshida, Pons, Maye, & Werker, 2010). Thus, acoustic implementation could plausibly contribute to a decline in sensitivity for allophonic distinctions in infants, aside from any other cues that they may draw from their proto-lexicon and their knowledge of complementary distributions.

In the current study, we investigated whether there are differences in the instantiation of allophonic and phonemic distinctions in infants' input. We tested the hypothesis that two sounds would be more acoustically distinct when produced phonemically than allophonically by comparing vowel tenseness and vowel nasality in American English and Quebec French. Vowel tenseness is phonemic in American English (*seat* [sit] versus sit [sɪt]) and allophonic in Quebec French (where vowels are laxed in closed syllables; e.g., *ville* [vɪl] 'city' versus *village* [vi.laʒ] 'town'; Rose & Wauquier-Gravelines, 2007; Walker, 1984). Vowel nasality is allophonic in American English (where vowels are nasalized in syllables closed by nasal consonants; e.g., *bet* [bɛt] versus *Ben* [bɛ̃n]; Krakow, 1993) and phonemic in Quebec French (*monde* [mõd] 'world' versus *mode* [mod] 'fashion').

We measured acoustic–phonetic correlates of tenseness and nasality in the speech of three groups of mothers. One group lived in the American Midwest and spoke only American English. A second group lived in Montreal and spoke primarily Quebec French. The third group consisted of bilinguals living in Montreal, who spoke both Quebec French and Quebec English frequently, from an early age, and without a marked foreign accent. Each mother was recorded while talking with their 11-month-old about sets of objects, which had been selected so that the object labels contained specific target vowels. This infant-directed corpus contained data on two pairs of vowels differing in tenseness ([i-ɪ], for both English and French, and either [e-ɛ] for French or [eɪ-ɛ] for English) and two pairs of vowels differing in nasality ([ɛ-ɛ̃] and either [a-ã] for French or [æ-æ̃] for English). For each mother and for each vowel pair, we assessed the separation between the vowel centroids along several perceptually-relevant acoustic correlates.

If the perception–production loop described above is at least partially accurate, then we predict that the separation between [i-ɪ] will be larger in English than in French, as will the separation between [e(ɪ)-ɛ]. In contrast, the separation between [ɛ-ɛ̃] should be larger in French than in English, and the same should occur for [a/æ -ã/æ̃]. Our predictions only operate *within* contrast type, and not across them. Our key prediction is that the acoustic implementation of a given distinction might be affected by whether it is phonemic or allophonic (through the perception–production loop described above), and thus no predictions can or need to be made with respect to other contrasts in the same system. (Additionally, such a direct comparison would also be empirically inadequate, as we will see below.)

1.2. Phonetic instantiation in bilingual speech

The present study elaborates on data from both monolinguals and bilinguals. The latter population contributes to our research question in two ways.

First, the speech of bilinguals constitutes a strong test of our hypothesis. As noted above, a difference in acoustic–perceptual distance depending on phonological status should follow from the structure of the phonological and lexical systems, and it would be reinforced by perception and production practice. Therefore, such a difference should be evident *even in the speech of the same person* regularly using two different language