Converging Evidence for Abstract Phonological Knowledge in Speech Processing

Anne Cutler (a.cutler@westernsydney.edu.au)
The MARCS Institute, and ARC Centre of Excellence for the Dynamics of Language,
Western Sydney University, Locked Bag 1797, Penrith NSW 2751, Australia;
Max Planck Institute for Psycholinguistics, P.O. Box 310, Nijmegen 6500 AH, The Netherlands.

Introduction

The perceptual processing of speech is a constant interplay of multiple competing albeit convergent processes: acoustic input vs. higher-level representations, universal mechanisms vs. language-specific, veridical traces of speech experience vs. construction and activation of abstract representations.

The present summary concerns the third of these issues. The ability to generalise across experience and to deal with resulting abstractions is the hallmark of human cognition, visible even in early infancy. In speech processing, abstract representations play a necessary role in both production and perception. New sorts of evidence are now informing our understanding of the breadth of this role. Two earlier and more detailed reviews of the role of abstraction in speech processing (Cutler, 2008; 2010) also embrace, respectively, evidence on the lexical representation of form versus meaning, and evidence on prosodic processing.

Evidence from the second-language lexicon

Learners of a second language (L2) endure persistent perceptual trouble (not fixable just by accruing experience) with L2 phonemic distinctions that their first language (L1) lacks. The classic explanation of this is that L1 phonology (here, abstract knowledge about which phonemic contrasts may be encountered in speech signals) captures the input.

This is not the full account of this difficulty, however. Abstract knowledge of a contrast's existence (from reading, or from teaching, e.g., that *light* and *write* are supposed to be different) influences the construction of phonological representations in the lexicon. In the L2 lexicon, these representations thus become distinct. Speech perception, however, still fails to deliver the discrimination this requires (Weber & Cutler, 2004; Cutler, Weber & Otake, 2006; Broersma & Cutler, 2008); both L2 sounds are perceived as (more or less good) realisations of a single phoneme. (That will usually be the L2 phoneme acoustically closest to the single L1 sound; for Japanese listeners hearing English [r/l], this is [l]). In *righteous* or *rightful*, the initial syllable will then actually activate *light*, not *right*. The second syllable will be needed in order to produce the desired lexical entry as the closest match to the input as a whole. The result is that word recognition in L2 is slower than it should be, because more competitor words are activated, and the competition from the spuriously activated ones is also more persistent (Broersma & Cutler, 2011; Cutler, 2015).

Evidence from talker adaptation

We adapt so rapidly to talkers we have never before heard by using existing knowledge to resolve phonetic ambiguity, and in consequence adjusting phoneme category boundaries for that specific talker (Norris, McQueen & Cutler, 2003; Eisner & McQueen, 2005). The adjustment generalises to words and phonetic contexts in which the phonemes in question have not previously been heard from the new talker (McQueen, Cutler & Norris, 2006). Thus the adaptation has concerned phonemic categories, not veridical traces of experience (Cutler, 2010). Episodic models of lexical storage and retrieval, in which stored traces of lexical experience are activated in proportion to their match to the current input, cannot cope with this generalisation result (Cutler, Eisner, McQueen & Norris, 2010), because the models are unable to assign the novel pronunciation instances uniquely to the phonemic category they should represent.

Evidence from cross-modal generalisation

Cross-modal priming is popular in psycholinguistics, not necessarily because it calls on representations abstracted across different modalities; it is just a robust and useful task. Interestingly, recourse to the supra-modal representation even informs priming across modalities when the target is the same articulatory event – hearing words facilitates later phonological processing from lipreading the same spoken words, compared with new words (van der Zande, Jesse & Cutler, 2014a). Notably, the lipreading here was facilitated whether or not the talker was the same one who had been heard in the priming phase; there was always an advantage for old words over new, but no effect of talker familiarity. Talker adaptation (as outlined in the section above) was likewise unaffected by visual information indicating another talker (van der Zande, Jesse & Cutler, 2014b). These results confirm that phonological representations in the lexicon are shared across auditory and visual processing, and also show that talker information is not transferred across modalities at the lexical level. The abstract representations are stronger than, or unaffected by, modality-specific experience.

Evidence from talker recognition

One of the best-known effects in talker recognition is that listeners find it easier to recognise talkers (pick them out from a set, as in a forensic lineup) when they are talking the listeners' native language. This turns out not to be due to a need to understand what is being said, because this native-language effect appears even with seven-month-old infants:

Dutch-learning infants at this age perk up when a new talker is added to a set of three female talkers uttering unrelated (adult-style) Dutch sentences, but do not notice when a new talker is added to a set speaking Italian, or a set speaking Japanese (Johnson, Westrek, Nazzi & Cutler, 2011).

At seven months, infants are acquiring the phonological structure of the language around them, but do not yet have a functional vocabulary that would allow them to understand such input. Thus the effect is here based on familiarity with the phonology in the one set of input but not in the others.

Analogously, adult listeners show equivalent efficiency with two phonologically comparable dialects of a language as opposed to a language with a differing phonology, and this works each way – native speakers of one of the dialects recognise talkers equally well in either dialect (but worse in the phonologically different language), while non-native listeners perform the talker recognition task equally badly in either dialect (but better in their phonologically different own tongue; Johnson, Bruggeman & Cutler, in press). Again, the phonological familiarity predicts the results.

Evidence from a lost language

Children adopted into another country lose all conscious knowledge of their first language and become essentially native speakers of a new language. But traces remain of the first, as many studies, with many languages, have shown. A recurring finding is that adoptees (in comparison to controls) show an accelerated trajectory of learning phonological structures found in the birth language but not in the current native tongue. In the largest such adoptee study so far, we replicated this for speech perception (Choi, Broersma & Cutler, 2017), and also found that the perceptual mastery transferred to speech production (Choi, Cutler & Broersma, 2017). This transfer, and a further generalisation of training on one phoneme contrast to other places of articulation, indicate that the observed benefit is based on abstract phonological representations. Most strikingly, the adoptee benefit was independent of age at adoption; infants adopted under the age of six months (before vocabulary building, or phoneme repertoire mastery, or talking) showed as much evidence of phonological retention as those adopted over the age of one. Thus abstract phonological knowledge is compiled and laid down even before six months of age, in preparation for the later stages of language acquisition.

Conclusion

Abstract phonological knowledge plays a role in all aspects of speech processing. This is true even of those processing realms which may seem to form natural sources of evidence for memory-based effects. Thus we can see that abstractions are involved in many kinds of processing where differences between talkers are at issue. Likewise, though phonological structures are language-specific and hence not inborn, whereby language acquisition needs speech input to set it going, it also appears that construction of abstract phonological generalisations across this input must form part of linguistic processing even in the earliest months of life.

References
Broersma, M. & Cutler, A. (2008). Phantom word recognition in L2. *System: An International Journal of Education Technology and Applied Linguistics, 36*, 22-34.

Broersma, M. & Cutler, A. (2011). Competition dynamics of second-language listening. *Quarterly Journal of Experimental Psychology, 64*, 74-95.

Choi, J., Broersma, M. & Cutler, A. (2017). Early phonology: Birth-language retention from the earliest months of international adoptees' life. *PNAS*, to appear.

Choi, J., Cutler, A. & Broersma, M. (2017). Early development of abstract language knowledge: Evidence from perception-production transfer of birth-language memory. *Royal Society Open Science, 4*, 160660.

Cutler, A. (2008). The abstract representations in speech processing. *Quarterly Journal of Experimental Psychology, 61*, 1601-1619.

Cutler, A. (2010). Abstraction-based efficiency in the lexicon. *Laboratory Phonology, 1*, 301-318.

Cutler, A. (2015). Representation of second language phonology. *Applied Psycholinguistics, 36*, 115-128.

Cutler, A., Eisner, F., McQueen, J.M. & Norris, D . (2010). How abstract phonemic categories are necessary for coping with speaker-related variation. In C. Fougeron, B. Kühnert, M.P. d'Imperio & N. Vallée (Eds.), *Papers in Laboratory Phonology 10*. Berlin: Mouton de Gruyter.

Cutler, A., Weber, A. & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics, 34*, 269-284.

Eisner, F. & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics*, 67, 224–238.

Johnson, E.K., Bruggeman, L. & Cutler, A. (in press). Abstraction and the (misnamed) Language Familiarity Effect. *Cognitive Science*.

Johnson, E.K., Westrek, E., Nazzi, T. & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science, 14*, 1002-1011.

McQueen, J.M., Cutler, A. & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science, 30*, 1113-1126.

Norris, D., McQueen, J.M. & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*, 204-238.

Weber, A. & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*, 1-25.

Zande, P. van der, Jesse, A. & Cutler, A. (2014a). Hearing words helps seeing words: A cross-modal word repetition effect. *Speech Communication, 59*, 31-43.

Zande, P. van der, Jesse, A. & Cutler, A. (2014b). Cross-speaker generalisation in two phoneme-level perceptual adaptation processes. *Journal of Phonetics, 43*, 38-46.