A GENERAL AUDITORY BIAS FOR HANDLING SPEAKER VARIABILITY IN SPEECH? EVIDENCE IN HUMANS AND SONGBIRDS

BUDDHAMAS KRIENGWATANA
Department of Psychology, University of Amsterdam,
Amsterdam, the Netherlands
bkrieng@alumni.uwo.ca

PAOLA ESCUDERO[1], ANNE H KERKOVEN[2], CAREL TEN CATE[2]
[1]*MARCS Institute, University of Western Sydney*
Sydney, Australia

[2]*Institute for Biology Leiden, Leiden University*
Leiden, the Netherlands
Paola.Escudero@westernsydney.edu.au, annekerkhoven@hotmail.com,
c.j.ten.cate@biology.leidenuniv.nl

1. Abstract

Different speakers produce the same speech sound differently, yet listeners are still able to reliably identify the speech sound. A compelling example of our ability to distinguish speech sounds despite enormous variability arising from speaker, gender, and age differences is in the case of vowels. Despite the large between-speaker variation within a vowel category and striking overlap between vowel categories, human adults, pre-linguistic infants, and even nonhuman animals are able to classify vowels of different speakers and genders. How is this achieved, and are they achieved in the same way by human adults, infants, and nonhuman animals?

Perceptual adjustments to accommodate for speaker differences in vowels may possibly be achieved pre-attentively via low level processing mechanisms. Combined with findings suggesting nonhuman animals also adjust for speaker differences, this raises an intriguing possibility that there is a tendency for the vertebrate auditory system to automatically accommodate for speaker differences in vowel production. If this is the case, then exposure to speaker-

variability in vowel production need not be necessary in order for listeners to compensate for speaker and gender differences.

The aim of this study was to compare the ability of humans and zebra finches to categorize vowels despite speaker variation in speech in order to test the hypothesis that accommodating speaker and gender differences in isolated vowels can be achieved without prior experience with speaker-related variability.

Using a behavioral Go/No-go task and identical stimuli, we compared Australian English adults' (naïve to Dutch) and zebra finches' (naïve to human speech) ability to categorize / ɪ/ and /ɛ/ vowels of an novel Dutch speaker after learning to discriminate those vowels from only one other speaker. The Go/No-go task requires subjects to make a response towards vowel stimuli assigned to one category (Go) and to inhibit responses toward vowel stimuli assigned to the other category (No-go). Experiments 1 and 2 presented vowels of two speakers interspersed or blocked, respectively. If experience with speaker variability in vowel production is necessary for successful normalization to occur, then we predicted that zebra finches and humans would not be able to discriminate the vowels of the second, new speaker.

Results demonstrate that categorization of vowels is possible without prior exposure to speaker-related variability in speech for zebra finches, and in non-native vowel categories for humans. This study is the first to provide evidence for what might be a species-shared auditory bias that may supersede speaker-related information during vowel categorization, although it is also possible that different perceptual mechanisms underlie performance in humans and songbirds. The role of experience with speaker-related variability may be to tune the auditory system to the most relevant acoustic parameters that define phonetic categories. Our results do not seem to be adequately explained by existing vowel normalization algorithms (e.g. formant ratios, exemplar-based models). Future investigations of alternative accounts of vowel normalization should incorporate the possibility of an auditory bias for disregarding between-speaker variability, and bear in mind that there are many similarities between humans and zebra finches in vocal production, characteristics of the acoustic vocal signal, auditory perception, and need for accurate perceptual categorization. Thus, it may not be so surprising that there are also parallels in perceptual mechanisms that allow both species to overcome the problem of separating variability associated with content of the signal from variability arising from the individual signaler.