

Commentary: Kriengwatana, B., Escudero, P. & ten Cate, C. Revisiting vocal perception in non-human animals: a review of vowel discrimination, speaker voice recognition, and speaker normalization (2015) *Front. Psychol.* 5:1543

Linda Polka, Ocke-Schwen Bohn and Daniel J Weiss

Journal Name:	Frontiers in Psychology
ISSN:	1664-1078
Article type:	General Commentary Article
First received on:	18 May 2015
Revised on:	18 Jun 2015
Frontiers website link:	www.frontiersin.org

Commentary: Kriengwatana, B., Escudero, P. & ten Cate, C. Revisiting vocal perception in non-human animals: a review of vowel discrimination, speaker voice recognition, and speaker normalization (2015) *Front. Psychol.* 5:1543

Linda Polka¹, Ocke-Schwen Bohn² & Daniel J. Weiss³

¹ School of Communication Sciences Disorders and Centre for Research on Brain, Language and Music, McGill University, Montreal, Canada

² Department of English and Center on Autobiographical Memory Research, Department of Psychology, Aarhus University, Denmark

³ Department of Psychology and Program in Linguistics, The Pennsylvania State University, University Park, USA

Correspondence: School of Communication Sciences Disorders and Centre for Research on Brain, Language and Music, 2001 McGill college avenue 8th floor, McGill University, Montreal, Canada E-mail: linda.polka@mcgill.ca

Keywords: comparative research, vowel perception asymmetries, Natural Referent Vowel framework, infants, nonhuman animals, task demands

Comparative research provides a unique window into our understanding of human vocal perception. We commend Kriengwatana, Escudero & ten Cate (KEtc) for providing a much-needed review of this diverse literature. Their appraisal of three research areas highlights conceptual and empirical gaps, while also pointing to fruitful directions for future research.

This commentary addresses the literature on asymmetries in vowel perception. In their review of this topic KEtC focus on vowel contrasts that have revealed directional asymmetries in infants and nonhuman animals. We offer some clarification with respect to these stimulus issues and highlight another aspect of this research landscape – the role of task demands – that must also guide future comparative investigations.

Vowel perception asymmetries – Stimulus issues

The authors present a detailed overview of studies that reveal directional asymmetries in vowel discrimination in infants and several nonhuman species. To date infant perceptual asymmetries can be accounted for within the NRV framework (Polka & Bohn, 2011). Although directional asymmetries are evident for vowel pairs tested with cats, vervet monkeys, birds, and macaques, the overall pattern of the asymmetries observed in these species is inconsistent with the predictions of the NRV model.¹ As KEtC note, contrast-specific comparisons are limited because very few contrasts have been tested with both animals and infants. However, the richest cross-species data set pertains to the /ε-æ/ contrast. For this contrast, infants show the asymmetry predicted by NRV (easier in the /ε/ to /æ/ direction) whereas cats, birds, and vervets show an asymmetry in the opposite direction. Macaques performed at ceiling in both directions.

These findings point to distinct vowel discrimination patterns in human infants and nonhuman animals. Surprisingly, KEtC dismiss these findings and question whether the /ε-æ/ asymmetry in infants is interpretable. They further suggest that we have not claimed that infant perception of /ε-æ/ supports the NRV framework; this is an incorrect representation of our work. In Polka & Bohn (1996), which subsequently led to the formulation of the NRV framework, we report and discuss the /ε-æ/ asymmetry and propose an account based on the location of these vowels in the vowel space (/æ/ is more extreme than /ε/). We further propose how this peripherality effect is acoustically grounded in Polka & Bohn (2011, page 474, paragraphs 6 & 7): “*The salience and stability of natural referent vowels is due to formant frequency convergence or focalization. ... Focalization is graded and gives rise to salience differentials across the vowel space.*” To clarify, in all of the infant and animal experiments on the /ε-æ/ contrast to date, focalization differences are clearly observed; i.e., F1 and F2 are spectrally closer in the more peripheral vowel /æ/ compared to the less peripheral /ε/. Accordingly, there is no basis for viewing research on /ε-æ/ as irrelevant to a discussion of comparative differences in vowel perception asymmetries.

This issue aside, we concur with KEtC that the current literature is sparse and inadequate for drawing firm conclusions regarding species-specificity with respect to vowel perception asymmetries.

Vowel perception asymmetries – task demands

In the current literature the tasks used to assess vowel discrimination asymmetries in infants and in other species are not comparable. These task discrepancies are not discussed by KEtC, yet they also severely limit the inferences that can be made. The animal studies cited by KEtC were conducted using psychophysical techniques designed to minimize cognitive resources making them ideal for comparing the peripheral sensory capacities of humans and nonhuman animals. In this work, a few subjects are extensively trained (with reinforcement) over many test sessions to discriminate a very small set of stimuli (one token per vowel), and memory demands are minimized by presenting the vowel stimuli with short inter-stimulus intervals (250 - 700 ms). This close temporal proximity allows the listener to access and compare acoustic details of the stimuli without encoding and retrieving information in a more enduring memory store. In contrast, the infant studies were conducted to understand how meaningful phonetic units are processed in more cognitively demanding tasks. Typically, a group of infants is tested using category-based discrimination tasks that involve stimulus variability (multiple tokens per vowel) and much less training, usually a single test session which may or may not involve reinforcement. Additionally, the temporal gaps between stimuli are longer (1000 -1500 ms), placing higher demands on memory.

The animal research has focused on perception of just-noticeable differences while the infant research has focused on just-meaningful differences. This distinction is critical in the context of vowel perception asymmetries. In the NRV framework asymmetries reveal vowel perceptual biases that emerge when perceivers are accessing phonetic units, not simply detecting acoustic differences. Thus, phonetic biases are predicted to surface in tasks that mirror at least some of the demands of natural speech processing (high memory demands, stimulus uncertainty). The tasks implemented in the current animal literature clearly do not tap this level of processing. The psychophysical tasks implemented with animals would likely yield ceiling effects in humans

which, interestingly, is the pattern found in macaques (Sinnott, 1989), a species with some ability to produce vowel-like sounds.

With respect to future research, we wholeheartedly agree with KEtC that testing human infants and nonhuman animals on the same vowel contrasts using comparable methods is required for drawing solid conclusions. More importantly, systematic manipulation of task demands is necessary to understand similarities and differences in the sensory, perceptual, and cognitive mechanisms across humans and non-human species. As highlighted by Weiss & Newport (2006) the perceptual and cognitive mechanisms that are fundamental to language acquisition cannot be adequately assessed with minimal stimulus variability/high training tasks. In the domain of vowel perception what is needed, as a minimum, are experiments that compare infants and other species in tasks that vary stimulus variability and memory load, and access to explicit training. Ideally, this would also involve comparing nonhuman primate species that vary in their capacity to produce vowel-like sounds.

Overall, when task differences are acknowledged, the current literature provides no compelling evidence that nonhuman animals show the kind of vowel perception biases that have been documented in humans. Despite the challenges, researchers are developing novel methods to assess perception in nonhuman animals across a wider range of processing demands. For example, several researchers have successfully measured spontaneous listening preferences in nonhuman primates (e.g. McDermott & Hauser, 2004; Watanabe & Nemoto, 1998). Understanding the evolution of language involves identifying which aspects of speech processing are shared with other animals and which are human-specific (Pinker & Jackendoff, 2005). To achieve this we must identify the dimensions of the speech signal that are accessed and also uncover the mechanisms that come into play when different species interact with spoken language.

ACKNOWLEDGMENTS

Support from Danmarks Grundforskningsfond (Danish National Research Foundation, grant DNRF93) to O-SB, from Natural Sciences & Engineering Research Council of Canada to LP, and from the National Institutes of Health (R01 HD067250) to DW is gratefully acknowledged.

REFERENCES

- McDermott, J. & Hauser, M. (2004). Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. *Cognition*, B11-21.
- Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, 95, 201–236.
- Polka, L. & Bohn, O.-S. (1996) A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America* 100, 577-592.

Polka, L. & Bohn, O.-S. (2011) Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics* 39, 467-478.

Sinnot, J. M. (1989) Detection and discrimination of synthetic English vowels by Old-World monkeys (*Ceropithecus Macac*) and humans. *Journal of the Acoustical Society of America*, 86, 557-565.

Watanabe, S. & Nemoto, M. (1998). Reinforcing properties of music in Java sparrows. *Behavioral Processes*, 43, 211-218.

Weiss, D.J. & Newport, E. L. (2006) Mechanisms underlying language acquisition: Benefits from a comparative approach, *Infancy*, 9(2), 241-257.

FOOTNOTES

¹ In KEtC's Figure 1, which plots asymmetries in discrimination of vervets from Sinnot (1989), the arrow going from /ʌ/ to /ɔ/ is in the wrong direction. The asymmetry described with respect to /u/-/ʊ/ is also reversed in the first paragraph on page three, but correctly displayed in Figure 1.