The Perceptual Learning of Time-Compressed Speech: A Comparison of Training Protocols with Different Levels of Difficulty

Yafit Gabay1,2, Avi Karni2,3,4 and Karen Banai5

1 Special Education Department and the 2 Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, Haifa, Israel
3 Sagol Department of Neurobiology and 4 Learning Disabilities Department, University of Haifa, Haifa, Israel
5 Department of Communications Sciences and Disorders, University of Haifa, Haifa, Israel

INTRODUCTION

Speech perception can be improved substantially with practice (perceptual learning) (Samuel & Kraljic, 2009).

Several approaches exist for skill acquisition including:

• Constant vs. variable training

• Errorless (Karni & Sagi, 1991) vs. desired level of difficulties (Bjork & Bjork, 2011).

It is still not clear which training yields the most perceptual learning of time-compressed speech.

GOAL

Determining which protocols’ features are important for the perceptual learning of time-compressed speech and its generalization

METHODS

Participants. 65 native Hebrew speakers.

Stimuli. 120 simple active sentences in Hebrew (Prior & Bentin, 2006).

Time-Compression. using a WSOLA algorithm (Verhelst & Roelands, 1993).

RESULTS

For trained tokens, lowest performance on the constant-high protocol. $t(24) = -5.38, p < .01$.

For new tokens – adaptive protocols are better than constant, $F(1, 60) = 5.83, p = .01$.

For untrained talker, a trend toward significance, $F(1, 48) = 3.57, p = .06$, implies the lowest performance in the constant high protocol.

CONCLUSIONS

Initial training on “easy items” (errorless learning) can facilitate the perceptual learning of speech - lowest performance when initial trials are difficult.

Speech-rate variability can support generalization - better performance in adaptive compared with constant protocols.

The differences between acoustic input and lexical representations are smaller in adaptive compared to constant protocols, thus enabling a more gradual adaptation.

Email: ygabay@edu.haifa.ac.il

References

1. Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. Psychology and the real world: Essays illustrating fundamental contributions to society, 56-64.
2. Prior, A., & Bentin, S. (2006). Differential integration efforts of mandatory and optional sentence constituents. Psychophysiology, 43(5), 440-449.
3. Karni, A., & Sagi, D. (1991). Where practice makes perfect: feature decomposition evidence for primary visual cortex plasticity. Proceedings of the National Academy of Sciences, 88(11), 4966-4970.
4. Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. Attention, Perception, & Psychophysics, 71(6), 1207-1218.
5. Verhelst, W., & Roelands, M. (1993). An overlap-add technique based on waveform similarity (WSOLA) for high-quality time-scale modification of speech. Paper presented at the Acoustics, Speech, and Signal Processing, 1993. ICASSP-93., 1993 IEEE International Conference on (Vol. 2, pp. 554-557). IEEE.

This research was supported by a grant from the National Institute of Psychobiology in Israel to KB.