THEME SESSION: SENSORY-COGNITIVE INTERACTIONS IN SPEECH COMPREHENSION

Bruce A. Schneider
University of Toronto Mississauga
bruce.schneider@utoronto.ca

Abstract

To comprehend speech, listeners must separate the information available in the speech signal from auditory interference produced by other concurrent sound sources (auditory scene analysis). Auditory scene analysis depends not only on a number of bottom-up sensory processes involved in signal extraction, but also on a number of knowledge-driven, top-town processes (for example, prior knowledge of the topic of conversation). Hence, how effective listeners will be with respect to the extraction of the relevant information will depend not only on their sensory acuity but also on their command of the language, and the effectiveness of the cognitive processes involved in these top-down processes. In this theme session we will discuss how age-related changes in sensory and cognitive processes alter the manner in which scene analysis is performed, and how the listener's command of the language, and the language itself affects a listener's ability to comprehend spoken language.

Psychophysics began as an attempt to link the physical dimensions of stimuli to their mental representations, with little or no concern about what the observer did with these representations or percepts. Cognitive science, on the other hand, typically has focused on how the perceptual objects constructed from these percepts were organized and interpreted in ways which permitted observers to navigate and comprehend their environment. More recently, evidence has been accruing that observers' expectations concerning their environment could affect basic sensory processes (dela Rosa et al., 2009; Mišić et al., 2010; Schneider et al., 2011) . Hence, there is accumulating evidence that psychophysicists need to consider how sensory and cognitive systems interact if they are to fully understand how sensory systems work (Schneider & Parker, 2010). In this theme session we will explore how sensory and cognitive processes interact in the domain of speech comprehension.

To comprehend speech, listeners must separate the information available in the speech signal from auditory interference produced by other concurrent sound sources (auditory scene analysis, Bregman, 1990). For example, in order to follow a conversation in a cocktail party, the listener has to perceptually segregate the target talker's voice from that of other talkers. In other words, the listener has to perceive the target talker as a distinct auditory "object" in an otherwise cluttered auditory scene. A number of studies have shown that the analysis of the auditory scene not only involves bottom-up signal-extraction processes, but also a number of knowledge-driven, top-town processes (for example, prior knowledge of the topic of conversation, see, for example, Ezzatian et al, 2011; Freyman et al., 2004). Hence, how effective listeners will be with respect to the extraction of the relevant information will depend not only on their sensory acuity but also on their command of the language, and the effectiveness of the cognitive processes involved in these top-down processes.

A number of studies have shown that a listener's ability to segregate a target voice in an auditory scene depends on the nature of the other sound sources in the environment. The first thing to note is that other concurrent sound sources will initiate activity along the basilar

membrane that will overlap that which is induced by the target voice. This overlapping activity will lead to peripheral or energetic masking of the target voice. Any sound source whose spectrum overlaps that of the target voice will produce energetic masking. However, there is a second type of masking in such situations, which is often referred to as informational masking, and is thought to act at more central levels. For example, a possible source of informational masking would be another person talking at the same time as the target talker. One way in which the competing voice could interfere with the comprehension of the target voice, would be if stream segregation (separating the target voice from the competitor's voice) was incomplete. Under such circumstances we might expect the information in the competing voice to interfere with the processing of the target voice at more central levels (i.e., with the processes involved in phonemic, semantic, and linguistic processing).

In this theme session we will show that the amount of both energetic and informational masking of speech by speech is affected by: 1) the status of the sensory systems; and 2) the language competence of the listener. Moreover, we will see that both types of factors interact in interesting ways, and that the pattern of this interaction helps to explain why listening in complex auditory scenes is so much more difficult for older adults than for younger adults, and for those operating in their second or third language.

Acknowledgements

The research in this article was supported by grants from the Natural Science and Engineering Research Council of Canada (RGPIN 9952) and the Canadian Institutes of Health Research (MOP 15359).

References

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sounds*. London, UK: The MIT Press.

de la Rosa, S., Gordon, M.S., & Schneider, B. A. (2009). Knowledge alters visual contrast sensitivity. *Attention, Perception & Psychophysics*, *71*, 451-462.

Ezzatian, P., Li, L., Pichora-Fuller, K., & Schneider, B. (2011). The effect of priming on release from informational masking is equivalent for younger and older adults. *Ear & Hearing, 32 (1)*, 84-96.

Freyman, R. L., Balakrishnan, U., & Helfer, K.S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *Journal of the Acoustical Society of America*, *115*, 2246-2256.

Mišić, B.V., Schneider, B.A., & McIntosh, A. R. (2010). Knowledge-drive contrast gain control is characterized by two distinct electrocortical markers. *Frontiers in Human Neuroscience*, *3*, 1-12.

Schneider, B. A. & Parker S. (2010). The evolution of psychophysics: From sensation to cognition and back again. In Bastianelli, A., Vidotto, G. (Eds.). *Fechner Day 2010 (pp. 7-10). Proceedings of the 26th Annual Meeting of the international Society for Psychophysics*, Padua, Italy: The international Society for Psychophysics.

Schneider, B. A., Parker, S., & Murphy, D. (2011). A model of top-down gain control in the auditory system. *Attention, Perception, & Psychophysics, 73(5)*, 1562-1578.