

This is an excerpt from a report on the workshop on multisensory integration at the University of Toronto, on May 9th and 10th, 2014, written by Kevin Connolly, Aaron Henry, Zoe Jenkin, and Andrew MacGregor, and available at:
http://networksensoryresearch.utoronto.ca/Events_%26_Discussion.html

4. Is Language Processing a Special Kind of Multisensory Integration?

Language processing involves multisensory integration. For instance, as Teresa Blankmeyer Burke pointed out, the process of speech reading involves visual sensitivity not just to areas around the speaker's mouth, where one takes articulation to occur, but also to other regions of the face. Drawing in part from her own experience with speech reading, Burke described how speech reading is impaired not only if the speaker has facial hair (obscuring fine mouth movements) but also if they are wearing sunglasses.

In both research in perceptual psychology and in philosophical discourse, language is often treated as a special case of perception. This is certainly for good reason—language has many features that are distinct from features of perception more generally. For example, our language-processing faculty gives us the ability to understand and generate compositional linguistic structures, which is arguably unique to the domain of language (as opposed to say, olfactory or tactile states, which might well have content but do not display the same sort of recombinable structure). Questions about reference and content may also have different answers when applied to linguistic perceptual representations as opposed to when they are applied to perceptual states more generally. Might the question of how multisensory integration works in the linguistic domain similarly receive a fundamentally different answer than it does in non-linguistic domains?

Casey O'Callaghan offered an argument that language processing is not as special in this respect as it is often made out to be. In his talk, O'Callaghan outlined six grades of multisensory integration that occur in various different types of perceptual processing, none of which are necessarily or even predominantly linguistic. When asked where speech

processing fits into his map of the multisensory landscape, he claimed that it could be thoroughly explained as a combination of the different types of multisensory integration that he had already described. On this view, there would be no need to posit a unique species of multisensory perception for language, either in terms of the processing mechanisms involved, in terms of the types of contents, or in terms of the phenomenal character of such multisensory linguistic perceptual states. If this is the case, the multisensory aspects of speech processing would be by no means uninteresting, but they would be better thought of as among a group of special phenomena (multisensory perceptions) than distinctly special on their own, at least in this respect.

Barry Smith provided some reasons one might think that language processing is a special kind of multisensory integration--different from standard cases involving audition but not speech. First, while in cases of non-linguistic audition, the perceptual object is sounds, it is unclear what the perceptual object of speech would be. Speech perception need not be the perception of sounds, but could be any number of things: sounds with meanings, meanings, a voice, the speaker, someone saying such-and-such, or some further option. Potentially, this makes speech special from cases of non-linguistic audition. Second, the brain treats hearing a human voice do something non-linguistic (such as groaning, laughing, or crying) differently from hearing a human voice speak. Smith cited research showing that immediately upon a subject's identification of electronically produced sounds as speech, there is a correlated transfer from the general auditory cortex to specialized areas in the language centers involved in the processing of speech. This seems to indicate, again, that speech is special, in that it is distinct from other kinds of auditory perception.

In her talk, Janet Werker presented an intermediate position on the question of whether language processing is special or not when compared to other types of multisensory processing. On the one hand, Werker asserted that language processing is special in that

many central features of language mastery, such as categorical distinctions between phonemes, aren't acquired by domain-general learning mechanisms like association, but instead draw upon a proprietary set of representations that prepare the infant for language acquisition. There are, Werker noted, significant constraints on the forms that can serve as a possible signal in a language. For example, a human being cannot learn a language composed wholly of mechanical sounds. Thus the possible forms that can serve as linguistic units are non-arbitrary, even if the meaning that we assign to these units is arbitrary. On the other hand, Werker denied that speech is unique in these respects. Apart from speech, sign language has also been shown to depend on categorical distinctions within the perceptual system to which human infants are sensitive from early in life and which prepare the infant for learning that language. Werker also referred to evidence suggesting that other natural sound signals, such as the sound of water, may qualify as "special" in the relevant sense. Thus she took a similar position to O'Callaghan, although for different reasons, that language processing is not entirely unique, nor is it entirely common. Instead, it's among a group of unique things.