GENERAL DISCUSSION OF SESSION 2: SEPARATION OF SIMULTANEOUS EVENTS

Chairman: Brian C.J. Moore

With regard to Bregman's paper, there seemed to be general agreement with Bregman's account of duplex perception, indicating that it was not unique to speech perception. The question remained, nevertheless, whether it gave any special insight into mechanisms of speech perception. It was suggested, that, if it were possible to investigate duplex perception in infants, it would be interesting to determine whether it occurred at a particular age, linked to the development of linguistic knowledge.

One objection to Bregman's account was that "transparency" is not in fact a commonly occurring property of sounds. Rather, one sound would only be audible "behind" another if the levels of the sounds were comparable. More normally, the more intense sound would mask the weaker one. Bregman's response was that the methods used by the auditory system to separate simultaneous sounds do not guarantee success. For example, temporal synchrony may be used to group sounds together, but this does not always work reliably in reverberant rooms. Hence the auditory system is prepared to make use of rules which work some of the time. In general, we operate in a redundant way, making use of several different methods to separate acoustic events. With regard to duplex perception, the key question is: are there ever situations where it is justified to use the same piece of information twice? The answer is "yes", and this explains why the auditory system is able to do this.

The question was raised as to whether the auditory system used methods of separating acoustic events which are specific to speech. The answer seems to be "yes". For example, we can achieve a reasonable degree of accuracy in identifying two simultaneous vowels which are synthesised with exactly the same envelope and fundamental frequency. In this case, there are no obvious acoustic cues which could be used to achieve the separation of the two sounds, and knowledge about the spectral shapes of the vowels must be used. Of course, the use of knowledge of this type is not restricted to speech; our knowledge of the typical spectral composition of notes played by musical instruments might allow a similar separation of simultaneous notes.

A question was raised about the relative importance for sound separation of temporal modulation patterns in different frequency bands, and of the tracking of individual harmonics over time. Some models for separation of simultaneous sounds assume a rather low degree of frequency selectivity, and make use mostly of the temporal pattern of modulation. Other models assume greater frequency selectivity and place a greater emphasis on tracking harmonics over
time. There was general agreement that both types of cues exist and are used by the auditory system.

It was pointed out that when listening to synthetic speech it is often possible to hear the intended speech sounds, while at the same time certain extraneous sounds are audible, not "belonging" to the speech. This seems to be an example of duplex perception. The relative levels of the different acoustic elements seem to be important in this effect, and in the demonstration of duplex perception in which the formant transitions only are presented to one ear. The formant transitions are both integrated with the steady-state portion of the sound in the opposite ear and are heard as a chirp. If the level of the transitions is reduced, however, then a point is reached when they are not heard as a separate sound, but nevertheless they still form an integrated percept with the sound in the opposite ear. It is as if the phonetic processor takes its "share" of the sound first, and the auditory processor takes what is left (if anything).