On intelligibility, acoustic glimpses, and glimpsing rate

Oded Ghitza

In a recent study, Ghitza and Greenberg (2009) measured the intelligibility (in terms of word errors) of time-compressed speech with periodic insertions of silence. Three classes of signal conditions were considered: (1) waveforms without time compression, (2) waveforms with time compression, with a compression ratio of 3, and (3) waveforms in class 2 with further distortion inflicted: the waveforms were blindly segmented into consecutive 40-ms fragments, each followed by a silent interval. The parameters varied were the length of the silent interval, from 20 to 160 ms. The speech intervals were kept the same for all conditions.

For the signals in class 1 word error rate is marginal (ca. 2%). For the signals in class 2 performance is poor (> 50% word error rate). Surprisingly, for the signals in class 3 performance *improved*, with a U-shaped performance curve. The lowest word error rate (i.e., highest intelligibility) occurred when the silence was 80 ms long, down to ca. 20% (the "optimal" point). As far as I know there is no model of speech perception that can account for this data. Ghitza and Greenberg suggested that these data are consistent with a new model (termed TEMPO) in which brain rhythms affect the ability to decode the speech signal.

Before outlining the current study two terms are coined. Let the speech interval (40-ms long in Ghitza and Greenberg) be termed "*acoustic glimpse*"; and let the inverse of one period of the resulting signal (i.e. the inverse of the duration of a speech interval followed by a silence interval) be termed "*glimpsing rate*" (for example, a 40-ms long acoustic glimpse followed by a silent interval of 60-ms – a combined duration of 100 ms – results in a waveform with a glimpsing rate of 10 Hz).

In the current study we extend the experiments of Ghitza and Greenberg (2009) by considering a wider range of parameters. For example, in Ghitza and Greenberg, the time-compression ratio and the length of the acoustic glimpse were fixed (3 and 40 ms, respectively). In the current experiment, we'll consider the three parameters: time-compression ratio, acoustic glimpse and glimpsing rate, as variables. The experimental corpus and the experimental procedure will be as in Ghitza and Greenberg. The ultimate objective of the study is to reveal interrelationships between the various brain rhythms hypothesized to be involved in the speech decoding process (in the theta, beta and gamma bands, in particular).

Ghitza, O. and Greenberg, S. (2009). "On the possible role of brain rhythms in speech perception: Intelligibility of time compressed speech with periodic and aperiodic insertions of silence." Phonetica 66:113–126. doi:10.1159/000208934