Spontaneous speech adaptations in challenging communicative conditions across the lifespan

Plenary speech by

Valerie Hazan and Outi Tuomainen
Dept of Speech, Hearing and Phonetic Sciences, UCL, UK

Most of our knowledge about the acoustic-phonetic characteristics of speech come from speech production studies that have analysed controlled materials such as read sentences produced in isolation in a quiet environment. In typical communicative situations, the speech that we produce is likely to differ from such norms: it will be spontaneous, produced with true communicative intent, in less than ideal acoustic environments and quite often in a multi-tasking situation. In such situations, speech can be highly dynamic as we ongoingly adapt the level of clarity of our speech according to the demands of the communicative conditions, as suggested by Lindblom in his Hyper-Hypo model of speech production (Lindblom, 1990).

In our work, our aim has been to analyse the type of acoustic-phonetic adaptations made by speakers to counter the effects of adverse environments such as occur when communicating in noise, in the presence of other voices or with a hearing loss. We have recorded speech in laboratory conditions but have modelled natural communication by using a problem-based task that is carried out between two speakers. This picture-based 'spot the difference' task, called diapix (van Engen et al., 2010) involves the pair of speakers having to find 12 differences between their two pictures without seeing their conversational partner's picture. The degree of ease or difficulty with which speakers can communicate can be controlled by adding a communication barrier (e.g. babble noise, simulated hearing loss) affecting one or both of the speakers while they carry out the task together. This leads the person leading the interaction to naturally make adaptations to their spontaneous speech, producing a 'clear speaking style' in order to maintain effective communication, just as would happen in natural interactions. As speakers are recorded individually in connected sound-treated booths and communicate via headphones, a 'clean' and high-quality speech signal is recorded for each speaker.

In three consecutive large-scale studies, we have investigated such speech adaptations in 40 young adults (Hazan and Baker, 2011), 96 children aged 9 to 15 years (e.g., Hazan et al., 2016) and now 57 older adults aged 65 to 85 years, with 26 further younger adult controls (Tuomainen et al., 2016). A linked study has examined adaptations in hearing-impaired children while communicating with their hearing and hearing-impaired peers (Granlund et al., 2015). Each of these projects has led to the creation of large speech corpora (LUCID, kidLUCID and the forthcoming elderLUCID) containing many hours of spontaneous speech interactions. The lengthy processing of these corpora involves manual or automatic orthographic transcription, automatic alignment, manual checking of these alignments and the use of Praat scripts to obtain acoustic-phonetic measures. These measures include suprasegmental measures of articulation rate, fundamental frequency mean and range, relative intensity (representing spectral tilt) and segmental measures of vowel space.

In this talk, I will review our findings across these three studies spanning a broad age range; I will also discuss the challenges involved in the analysis of large spontaneous speech corpora. Our study with young adults showed that the adaptations that individual speakers made were, to an extent, dependent on the type of interference that was affecting their interlocutor (babble noise or vocoded speech), even though the speakers that we were analysing were not directly hearing the interference. This suggests that speakers used the direct or indirect feedback from their interlocutors during the interaction to attune their adaptations. Our study with children showed that they too made adaptations under similar conditions, although they had a tendency to use a strategy of increasing vocal effort (as shown by strong correlations between increases in fundamental frequency and decreases in spectral tilt) rather than using more varied strategies favoured by adults. Our ongoing study with older adults is showing a similar trend: older adults with age-related hearing loss tended to increase vocal effort to counter the effects of adverse conditions (again as shown by correlations between spectral tilt and fundamental frequency changes) while older adults with normal hearing thresholds and younger adults did not show this tendency.

In conclusion, investigating spontaneous speech in interaction in challenging communicative conditions can lead to a better understanding of the strategies used by speakers to maintain effective communication and of the impact of age and talker sex on such strategies. Despite the many challenges involved in the recording and analysis of spontaneous speech, such approaches will hopefully lead to a step forward in our knowledge of processes involved in speech communication.

References

Granlund, S., Hazan, V. L., & Mahon, M. (2015). Do children enhance phonetic contrasts in speech directed to a hearing-impaired peer? *Proceedings of the 18th International Congress of Phonetic Sciences*.

Hazan, V. L. & Baker, R. (2011). Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *Journal of the Acoustical Society of America* 130(4), 2139–2152

Hazan, V., Tuomainen, O. & Pettinato, M. (2016). Suprasegmental characteristics of spontaneous speech produced in good and challenging communicative conditions by talkers aged 9 to 14 years old. *Journal of Speech, Hearing and Language Research* 59, S1596–S1607.

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In Hardcastle, W.J. & Marchal, A. (eds) *Speech Production and Speech Modelling*, The Netherlands: Kluwer Academic, 403–439.

Tuomainen, O. & Hazan, V. (2016). Articulation rate in adverse listening conditions in younger and older adults. *Proceedings of Interspeech 2016*, 8–12 September 2016, San Francisco, USA.

Van Engen, K. J., Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., & Bradlow, A. R. (2010). The Wildcat Corpus of Native and Foreign-Accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech* 53, 510–540.