# The role of durational changes in the Lombard speech advantage

*Julián Villegas*[1,2], *Martin Cooke*[1,2], *and Catherine Mayo*[3]

[1]Ikerbasque (Basque Science Foundation), Spain
[2]Language and Speech Laboratory, Universidad del Pais Vasco, Spain
[3]Centre for Speech Technology Research, University of Edinburgh, UK
j.villegas@laslab.org,m.cooke@ikerbasque.org,catherin@inf.ed.ac.uk

## Abstract

Speech produced in the presence of noise—Lombard speech (LS)—has been found to be more intelligible than 'normal' speech when presented in equivalent amounts of noise [1, 2]. However, the origin of the LS advantage remains unclear. Part of the benefit appears to stem from spectral changes in LS which shift energy into the 0.6–3 kHz region (See Fig. 1) where it better escapes energetic masking by speech-shaped noise. Other parameters which show changes in LS include $f_0$ and duration. Lu & Cooke [3] modified the mean $f_0$ and spectrum (both independently and jointly) of normal speech, demonstrating a clear advantage of spectral modification but no effect of $f_0$.

The current study extended [3] in two directions. First, durational modifications to reflect differences between normal and LS were included. Second, as well as global (per utterance) changes, local (per frame) modifications were applied. Four male and four female talkers produced simple sentences containing spoken letter and number keywords in quiet and in intense speech-shaped noise (96 dB SPL). A perception experiment explored global vs. local modifications of spectral and durational parameters applied independently.

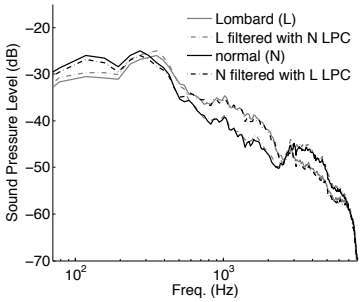



Figure 1: LTAS of the combined sentences used in the global spectral modifications.

Taken across all talkers, durational changes to normal speech produced no intelligibility benefit, whether applied globally (i.e., linear stretching or compression) or locally (using dynamic time warping to align normal and Lombard frames), spectral changes based on global or local modification were equally beneficial (See Fig. 2). However, when sentences were partitioned in two groups according to their speech rate in noise (See Fig. 3), some effect of durational modifications was observed: normal speech modified globally to the faster speech rate group was significantly less intelligible than unmodified speech, while conversely for the group with slower LS a small intelligibility benefit was present.

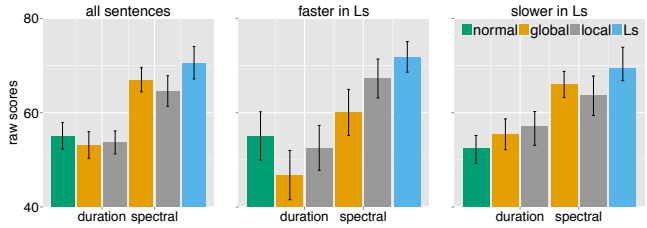
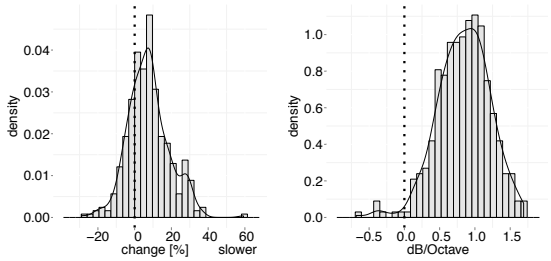



Figure 2: Raw intelligibility scores. Error bars correspond to the 95 % confidence interval.

These findings suggest that durational differences between normal and LS can affect intelligibility, but the benefit or otherwise depends on individual differences in speech rate.




(a) duration ($M = 7.58$) (b) spectral tilt ($M = 0.84$)

Figure 3: Differences between LS relative to normal.

**Acknowledgements.** The authors thank EU Future and Emerging Technology (FET–OPEN) Project LISTA (The Listening Talker) and staff at the Centre for Speech Technology Research of the University of Edinburgh for their collaboration running the experiment.

[1] J. J. Dreher and J. O'Neill, "Effects of ambient noise on speaker intelligibility for words and phrases," *J. Acoust. Soc. Am.*, vol. 29, no. 12, pp. 1320–1323, 1957.

[2] W. V. Summers, D. Pisoni, R. Bernacki, R. Pedlow, and M. Stokes, "Effects of noise on speech production: Acoustic and perceptual analyses," *J. Acoust. Soc. Am.*, vol. 84, no. 3, pp. 917–928, 1988.

[3] Y. Lu and M. Cooke, "The contribution of changes in f0 and spectral tilt to increased intelligibility of speech produced in noise," *Speech Comm.*, vol. 51, pp. 1253–1262, 2009.