Comparing the rhythmic properties of plain and Lombard speech

Hans Rutger Bosker[1,2] and Martin Cooke[3,4]

[1]Max Planck Institute for Psycholinguistics, PO Box 310, 6500 AH, Nijmegen, Netherlands
[2]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands
[3]University of the Basque Country, Universidad del País Vasco/EHU, Spain
[4]Ikerbasque (Basque Science Foundation), Bilbao, Spain

HansRutger.Bosker@mpi.nl

Speech is an inherently rhythmic signal. Even though speech is by no means strictly periodic, energy patterns in speech are constrained by the physiological dynamics of the lips, jaw, and tongue. As such, energy fluctuations in speech typically occur within the 2-20 Hz range.

This rhythmicity in speech has been suggested to play a central role in comprehension, facilitating the processing of the signal and aiding recognition memory. This study investigated whether speakers actually produce more rhythmic speech (i.e., more regular alternations between high and low amplitude intervals) in acoustically challenging conditions (e.g., background noise), thus facilitating comprehension.

Four different corpora were analyzed (with varying sample sizes and talkers), each including plain speech (sentences produced in quiet) and matched Lombard speech (same sentences produced in noise). Each sentence was first normalized in amplitude by RMS scaling, thus avoiding intensity confounds. The envelope of the normalized signal was then submitted to a Fast Fourier Transform (FFT), resulting in the modulation spectrum of that particular sentence. Comparing the average modulation spectra of plain and Lombard speech revealed greater power in Lombard speech in the *delta* band (1-3 Hz), across all four corpora, which could not be attributed to decreases in speech rate.

These findings suggest that speakers produce more rhythmic speech, particularly in the 1-3 Hz range, when talking in noise (vs. in quiet). Results are discussed in terms of the functional role of rhythmicity in dialogue and potential underlying neurocognitive mechanisms (e.g., neural oscillatory dynamics).