PROCEEDINGS of the
23rd International Congress on Acoustics
9 to 13 September 2019 in Aachen, Germany

Listeners' speech rate preferences in stationary and modulated maskers

Olympia SIMANTIRAKI[(1)], Martin COOKE[(2)]

[(1)]University of the Basque Country, Spain, olympia.simantiraki@ehu.eus
[(2)]Ikerbasque (Basque Science Foundation), Spain, m.cooke@ikerbasque.org

Abstract

Durational changes to speech can improve intelligibility in noise, but what speech rate would a listener choose in such conditions if allowed to do so? We introduce a technique which enables listeners to adjust arbitrary speech properties in real time while performing a speech-in-noise task, and apply it to the question of preferred speech rate. Native listeners adjusted speech rate while identifying word sequences in quiet, stationary noise (3 SNRs) and modulated noise (5 modulation rates). Listeners preferred faster rates in quiet, increasingly slower rates in stationary noise as SNR decreased, and increasingly slower rates in modulated noise as envelope modulation rate increased. These findings complement those obtained using traditional intelligibility measures. We speculate that for tasks where intelligibility is near to ceiling, any consistent listener preferences might be used as an indicator of optimal listening effort and applied in the design of speech enrichment algorithms. The listening preference technique can be adapted readily to many scenarios such as different masking conditions (e.g. competing talkers) and applied to other speech features such as F0 and spectral tilt.

Keywords: Speech rate, Speech-in-noise, Speech recognition task

1 INTRODUCTION

Listening to speech under imperfect conditions can reduce intelligibility (1) and lead to greater listening effort (2). Intelligibility enhancement in challenging conditions has been widely investigated, with speech modification algorithms able to achieve significant improvements (3). Recently, the goal of reducing listening effort has grown in prominence. Several studies have focused on listening effort by drawing comparisons amongst different listener groups (4, 5), maskers (6, 7), speech types (8), synthetic speech algorithms (9) and linguistic aspects of the speech signal (10). A number of objective and subjective measures have been used for quantifying listening effort (2, 11, 12). However, how specific factors such as speech rate impact on listening effort has received little attention (13). An alternative approach to extracting supra-intelligibility effects of speech is to allow listeners to control some aspect of speech during an intelligibility-based task. A similar approach was used in (14) to study listener speech rate preferences as a function of age and passage comprehension difficulty in quiet. The current study extends the listener preference approach to elicit preferred speech rates in a range of adverse listening conditions.

2 METHODS

Listeners were asked to change speech rate in real-time using up/down keys while listening to word sequences. The task was explained as akin to choosing an appropriate volume for a television: too quiet makes comprehension difficult, while too loud a setting leads to discomfort. Listeners were instructed to choose a speech rate that allowed them to recognise as many words as possible. Each trial consisted of an adjustment phase followed by a test phase in which their intelligibility was evaluated via a word identification task. The experiment was divided into 9 blocks by condition: quiet and 8 additive noise masking conditions: speech-shape noise (SSN) at SNRs of 0, +6 and +12 dB, and speech modulated noise (SMN) for 5 envelope modulation rates, mixed with speech at +6 dB SNR. Each block contained 22 trials. Eighteen native Spanish listeners were recruited and for each one, the experiment lasted around two hours with a short break between each block.

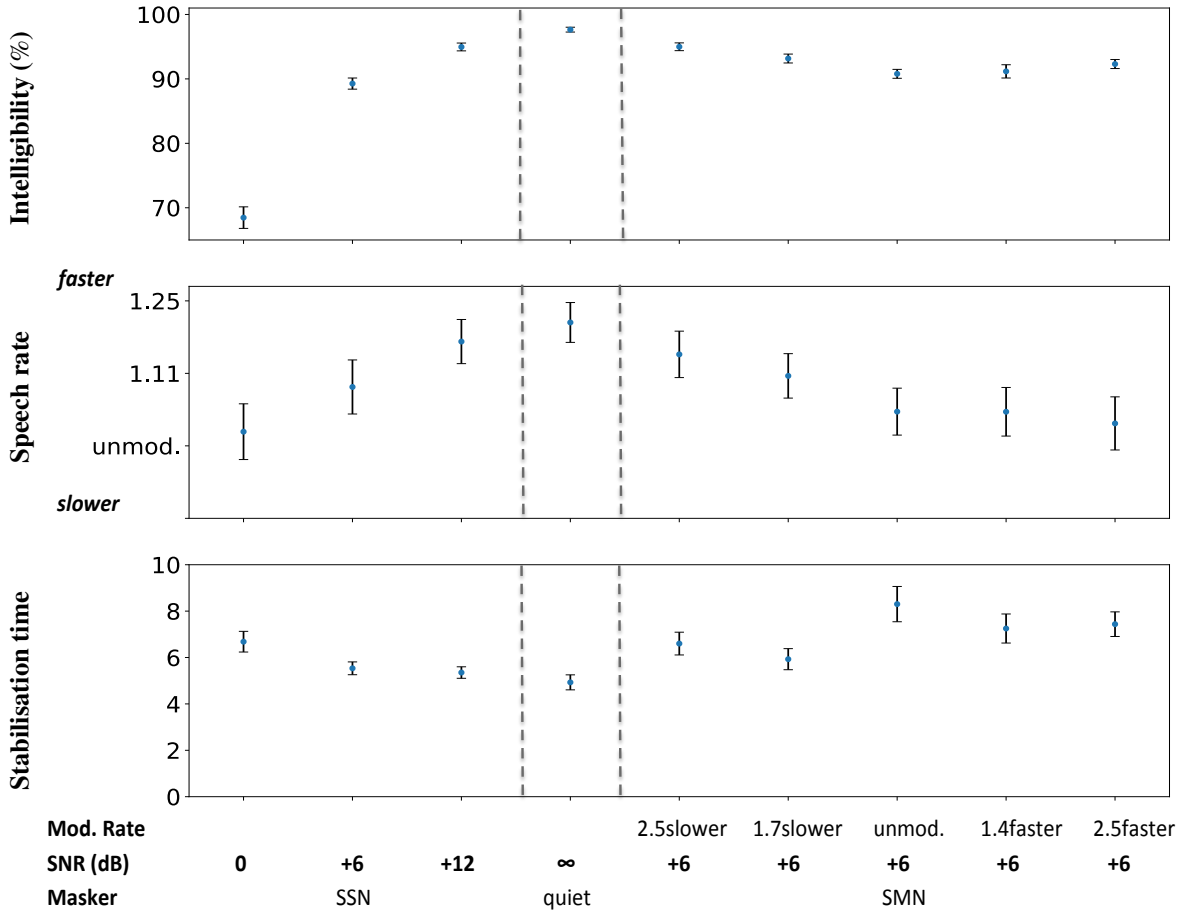

Figure 1. *Upper plot*: Mean of the percentage of correct words recalled. *Middle plot*: Mean of the preferred speech rate. Speech rate values correspond to the times slower/faster with regard to the original(*unmod.*). *Lower plot*: Mean time in seconds that listeners needed for stabilising their preference. Error bars represent ±SE.

3 SUMMARY OF THE FINDINGS

Listeners preferred slower speech rates for more challenging conditions (Fig. 1 *upper* & *middle*). For the SSN masker, the preferred rate became slower as SNR decreased ($p < .05$). This finding is in line with preferred rates in studies on speech complexity (14) and degraded (reverberant) listening conditions (15). For the SMN masker, listeners needed more time to express their listening preference compared to the quiet condition ($p < .01$). It may be that the modulated nature of the masker interferes with that of the target speech, causing attentional distraction or making it hard to predict when to listen. In all conditions the preferred speech rate was faster compared to the speech rate of the original recording, in agreement with (14). We speculate that stabilisation time may be an indicator of listening effort since different times are required to find the value that maximizes intelligibility in different conditions. Indeed intelligibility scores are not correlated with stabilization time ($r = -0.14, p = .09$).

ACKNOWLEDGEMENTS

Olympia Simantiraki was funded by the European Commission under the Marie Curie European Training Network ENRICH (675324).

REFERENCES

[1] Assmann, P.; Summerfield, Q. The Perception of Speech Under Adverse Conditions, Speech Processing in the Auditory System, Springer New York, Vol 18, 2004, pp 231-308.

[2] McGarrigle, R.; Munro, K.J.; Dawes, P.; Steward, A.J.; Moore, D.R.; Barry, J.G.; Amitay, S. Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'whitepaper', International Journal of Audiology, Vol 53, 2014, pp 433-445.

[3] Cooke, M.; Mayo, C.; Valentini-Botinhao, C. Intelligibility-enhancing speech modifications: the Hurricane Challenge, Interspeech 2013, 2013, pp 3552–3556.

[4] Borghini, G.; Hazan, V. Listening Effort During Sentence Processing Is Increased for Non-native Listeners: A Pupillometry Study, Frontiers in Neuroscience, Vol 12, 2018, pp 152.

[5] Ohlenforst, B.; Zekveld, A.A.; Lunner, T.; Wendt, D.; Naylor, G.; Wang, Y.; Versfeld, N.J.; Kramer, S.E. Impact of Stimulus-Related Factors and Hearing Impairment on Listening Effort as Indicated by Pupil Dilation, Hearing Research, Vol 351, 2017, pp 68-79.

[6] Koelewijn, T.; Zekveld, A.A.; Festen, J.M.; Kramer, S.E. Pupil Dilation Uncovers Extra Listening Effort in the Presence of a Single-Talker Masker, Ear and Hearing, Vol 33, 2012, pp 291-300.

[7] Zekveld A.A.; Kramer, S.E.; Festen, J.M. Cognitive Load During Speech Perception in Noise: The Influence of Age, Hearing Loss, and Cognition on the Pupil Response, Ear and Hearing, Vol 32, 2011, pp 498-510.

[8] Simantiraki, O.; Cooke, M.; King, S. Impact of Different Speech Types on Listening Effort, Proc. Interspeech 2018, 2018, pp 2267–2271.

[9] Govender, A.; King, S. Using Pupillometry to Measure the Cognitive Load of Synthetic Speech, Proc. Interspeech 2018, 2018, pp 2838–2842.

[10] Wendt, D.; Dau, T.; Hjortkjær, J. Impact of Background Noise and Sentence Complexity on Processing Demands during Sentence Comprehension, Frontiers in Psychology, Vol 7, 2016, pp 345.

[11] Ohlenforst, B.; Zekveld, A.A.; Jansma, E.P.; Wang, Y.; Naylor, G.; Lorens, A.; Lunner, T.; Kramer, S.E. Effects of Hearing Impairment and Hearing Aid Amplification on Listening Effort: A Systematic Review, Ear and Hearing, Vol 38, 2017, pp 267-281.

[12] Wu, Y.H.; Stangl, E.; Zhang, X.; Perkins, J.; Eilers, E. Psychometric Functions of Dual-Task Paradigms for Measuring Listening Effort, Ear and Hearing, Vol 37, 2016, pp 660-670.

[13] Müller, J.A.; Wendt, D.; Kollmeier, B.; Debener, S.; Brand, T. Effect of Speech Rate on Neural Tracking of Speech, Frontiers in Psychology, Vol 10, 2019, pp 449.

[14] Wingfield, A.; Ducharme, J.L. Effects of Age and Passage Difficulty on Listening-Rate Preferences for Time-Altered Speech, The Journals of Gerontology: Series B, Vol 54, 1999, pp 199–202.

[15] Moore, R.E.; Adams, E.M.; Dagenais, P.A.; Caffee, C. Effects of reverberation and filtering on speech rate judgment, International Journal of Audiology, Vol 46, 2007, pp 154–160.