4pSC21

Effects of noise, native language, age, and speaker gender on intelligibility in a large corpus of read speech

Richard Wright
U. Washington Linguistics

Benjamin V. Tucker
U. Alberta Linguistics

Matthew C. Kelley
U. Washington Linguistics

Marina Oganyan
U. Washington Linguistics

INTRODUCTION

We examine speech intelligibility in noise and the effect of the following factors:

- Speaker gender
- Listener native language
- Listener gender
- Listener age

RESEARCH QUESTION

Does a large dataset of speech intelligibility teach us anything new about speech perception?

METHODS

- 720 IEEE ("Harvard") sentences (Panfili et al., 2017)
- Native English Speakers; from WA, OR, and ID
  10 women, 10 men

Masking Noise

- Corpus-shaped noise (steady)
- SNRs: -2, 0, +2 dB

Listeners

1520 English-speaking adults

- 913 L1, 607 L2
- 997 female, 493 male, 30 other
- Age range: 16-70
  (mean: 21, standard deviation: 4.29)
- UAlberta and UW undergraduate students

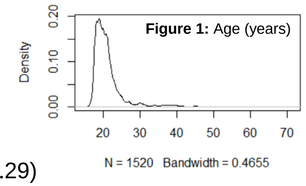

Figure 1: Age (years)

Task

- Online perception task (Nov. 2020 – Apr. 2022)
- Participants were asked to wear headphones
- 120 randomly selected test sentences
- Listeners typed each sentence to the best of their ability

Scoring

- Levenstein Distance across the entire sentence (Sohoglu & Davis, 2016)
- Jaro Distance across the entire sentence (Van der Loo, 2014)
- Calculated in R following Bosker (2021)

RESULTS

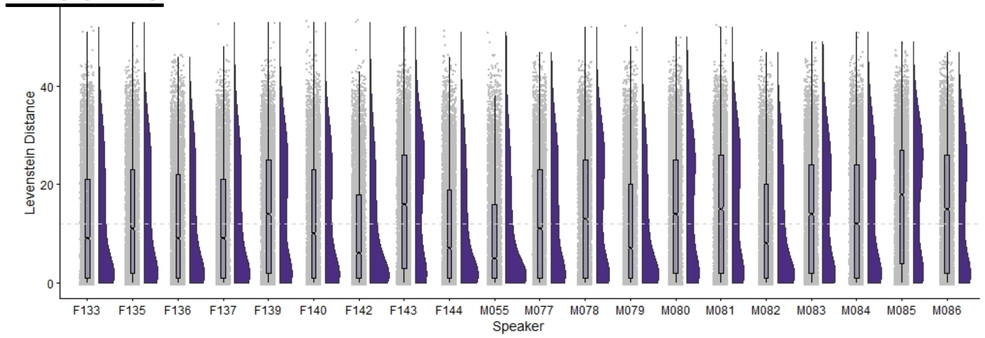

Figure 2. Raincloud plot of the Levenstein Distance split by speaker

- *Apparent gender effect (women > men)*
- Add a random effect for talker → *speaker gender effect goes away (cf. Fig. 2)*

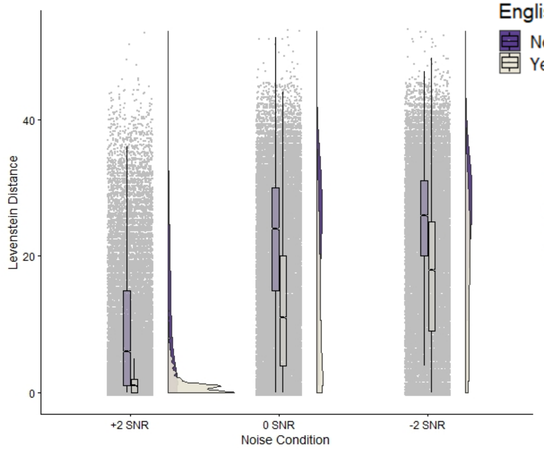

Figure 3: Levenstein distance split by native language

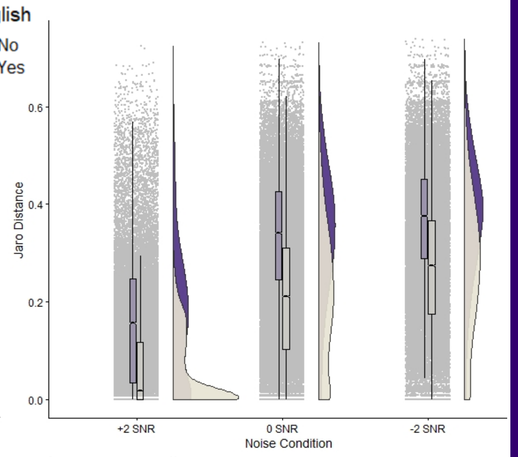

Figure 4: Jaro distance

- As expected, interaction with noise and native language for both measures of intelligibility
- *Native language difference in performance increases at poorer SNRs*

MODEL DETAILS

- Model fitting with lme4 (Bates et al., 2022) in R
- All talkers included, mixed model with random intercepts for sentence & listener
- `LD ~ NoiseCondition * EnglishFirstLanguage + ListenerGender + ListenerAge + SpeakerGender + (1|Speaker_ID) +(1|Listener)+(1|Item)`
- Model including Speaker RE preferred ($\chi^2(1)=9194$, $p<0.001$)

DISCUSSION

- This replicates findings that L2 listeners have greater trouble in noise than L1 listeners
- Suggests that any gender difference is driven by individual variability (i.e., a few unintelligible men) rather than a true group difference
- This replicates and extends previous findings showing **a lack of a gender** effect (McCloy et al., 2018)
  - Previous findings that women are more intelligible than men are likely not reliable
- Online testing allows for more efficient sampling
- A large population provides the statistical power to ask subtle questions (even with large variability in headphones)

FUTURE WORK

- More refined analysis of the available data

REFERENCES

Panfili, L. M., Haywood, J., McCloy, D. R., Souza, P. E., and Wright, R. A. (2017). The UW/NU Corpus, Version 2.0. https://depts.washington.edu/phonlab/projects/uwnu.php

McCloy D, Panfili L, John C, Winn M, and Wright R (2018). Gender, the individual, and intelligibility. Poster presented at the 176th Meeting of the Acoustical Society of America, Victoria, BC.

van der Loo, M. P. (2014). The stringdist Package for Approximate String Matching. *R Journal*, *6*(1), 111–122.

Sohoglu, E., & Davis, M. H. (2016). Perceptual learning of degraded speech by minimizing prediction error. *Proceedings of the National Academy of Sciences*, *113*(12), E1747–E1756. https://doi.org/10.1073/pnas.1523266113

Bosker, H. R. (2021). Using fuzzy string matching for automated assessment of listener transcripts in speech intelligibility studies. *Behavior Research Methods*. https://doi.org/10.3758/s13428-021-01542-4

ACKNOWLEDGMENTS

This work was supported by NIH NIDCD R01 DC006014

182nd meeting of the Acoustical Society of America (Denver, CO) May 2022