Canadian Audiologist

The Official Publication of the Canadian Academy of Audiology

Two Different Ways Familiar Voice Information Can Be Used

Published November 18th, 2018

Ingrid Johnsrude
Emma Holmes
Jon Deeks



Given the introductory paragraphs of most research papers on speech-in-noise understanding, one could be forgiven for thinking that individuals with hearing loss are hipper and better dressed than the rest of us – they seem to spend an awful lot of time at cocktail parties! The cocktail party is a clichéd example of the difficulties that such individuals face when trying to follow a conversation in the presence of competing talkers, but it is a valid one for two reasons. Trying to understand someone in the presence of competing talkers IS challenging, particularly when hearing loss is present. Second, conversational partners at cocktail parties can be highly familiar, or strangers, or anything in between. We are interested in how familiarity with a voice helps listeners to understand speech better when it is spoken by a friend or family member in the presence of competing speech.

Previous work[1–5] demonstrates that we are better at understanding what someone is saying if they are a spouse or close friend than if they are a stranger. Our previous work1 on intelligibility of spouse voices indicates that the benefit is quite large; equivalent to about 6–9 dB. This study used highly artificial materials that controlled the set of spoken words, controlled for context and meaning, and disrupted prosody: it probably underestimates the real benefit derived from naturally conversing with a friend or family member on real-world topics. Nevertheless, our results1 indicate that, over time, we learn something about the voices of the people we frequently talk to, which helps us to better understand the words they say.

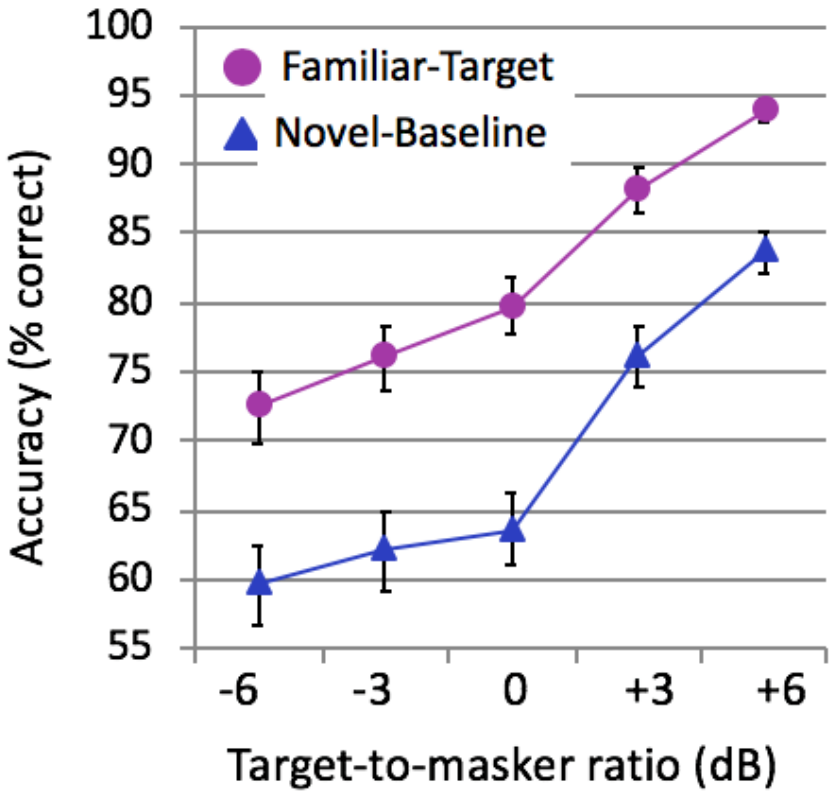


Figure 1. Accuracy at identifying words in spoken sentences when they are in a familiar voice, masked by an unfamiliar voice (purple circles) or when they are in an unfamiliar voice, masked by a different unfamiliar voice, as a function of target-to-masker ratio. The large effect of condition cannot be explained by acoustic differences, since the two conditions are identical over the group: voices were counterbalanced such that familiar voices also served as unfamiliar voices (for other listeners).

Of course, familiarity with a voice helps in another way as well – if a good friend phones you (to invite you to that cocktail party!), you know who they are within a couple of words. Their voice is like an "auditory face.[6]" In a recently published study[7] we asked: What features of a voice allow you to recognize a friend, and are they the same or different from the features that you use to help understand what they are saying? When we look at someone's face, the parts of the face we focus on to recognize who they are may differ from the parts we focus on to tell what mood they are in. Something similar might happen when we listen to voices.



We used software (Praat) to artificially manipulate two properties of recorded voices: voice pitch and the acoustic correlate of vocal tract length (formant spacing). We tested individuals listening both to original, and to manipulated, versions of the voices of a friend and two strangers (other participants' friends). For each voice sample, we asked participants if they could recognize the voice as their friend. We also, as in previous studies, compared the intelligibility of familiar and stranger voices when mixed with the voice of a competing talker (a stranger). On some trials, all of the voices were unmanipulated; on other trials the voices were manipulated either by shifting the fundamental frequency, or by shifting the formant spacing, or both. On any trial, all presented voices were manipulated the same way.

In manipulating the formant spacing of a familiar voice, we found that listeners were unable to recognize that the manipulated voice was their friend's voice. However, even though this manipulated voice was unrecognizable, people were still able to understand words spoken in this manipulated voice better than they could understand the same words spoken by a stranger.

Our findings demonstrate that we pick out different information from a voice, depending on whether we're simply trying to recognize the voice as a friend, spouse or parent on the phone, or whether we're trying to understand the words they're saying. This is relevant to the long-standing question of whether we separate what someone is saying (the speech content) from who is saying it (the identity of the person). The finding that speech is easier to understand when it is spoken by someone familiar demonstrates that we don't process the content of speech entirely separately from the identity of the person who is talking. Nevertheless, these new results show that we do process these characteristics somewhat differently – we must focus on different voice characteristics when we try to understand the content of words compared to when we try to recognize the identity of the person who is speaking.



For most of us, conversation is typically with people we know well. This is even more true of older people, who at the same time often experience hearing impairment. Understanding the acoustic properties of familiar voices that enhance speech intelligibility will enable engineers to ensure that hearing aids and cochlear implants preserve key vocal features of familiar voices, to maximize communication success and minimize effort in those with hearing impairment.

The next challenge is to understand how people become familiar with voices over time. How much experience with someone's voice is required for us to better understand what they're saying? Is this more or less than the length of experience we need to recognize someone from their voice? What experience is required to become familiar with a voice? We become familiar with radio hosts' voices, even if we have never seen them in person – but does this depend on us attributing an identity to them? If we can work out how people become familiar with a voice, we should be able to efficiently familiarize people with voices, and this might be useful for improving our ability to communicate in everyday situations.

Acknowledgements

Support for the research provided by that Natural Sciences and Engineering Research Council, the Canadian Institutes for Health Research, and by Western University's Canada First Excellence Research Fund BrainsCAN program.

References

1. Johnsrude IS. et al. Swinging at a cocktail party: voice familiarity aids speech perception in the presence of a competing voice. Psychol Sci 2013;24:1995–2004.
2. Nygaard LC and Pisoni DB. Talker-specific learning in speech perception. Percept Psychophy 1998;60:355–76.
3. Souza PE, Gehani N, Wright R, and McCloy D. The advantage of knowing the talker. J Am Acad Audiol 2013;24:689–700.
4. Yonan CA and Sommers MS. The effects of talker familiarity on spoken word identification in younger and older listeners. Psychol Aging 2000;15:88–99.

5. Newman RS and Evers S. The effect of talker familiarity on stream segregation. J Phonet 2007;35:85–103.
6. Belin P, Fecteau S, and Bédard C. Thinking the voice: neural correlates of voice perception. Trends Cog Sci 2004;8:129–35.
7. Holmes E, Domingo Y, and Johnsrude I. Familiar voices are more intelligible, even if they are not recognized as familiar. Psychol Sci 2018;29, 1575–83.