**Naturalness paper – 10 Important Theory Papers on Voice Perception**

(Belin, Bestelmeyer, Latinus, & Watson, 2011; Belin, Fecteau, & Bedard, 2004; Bhaya-Grossman & Chang, 2022; Deen, Koldewyn, Kanwisher, & Saxe, 2015; Frühholz & Schweinberger, 2021; Fruhholz, Trost, & Kotz, 2016; Gainotti, 2011; Hall, Horgan, & Murphy, 2019; Kreiman & Sidtis, 2011; Lavan & McGettigan, 2023; Moerel, De Martino, & Formisano, 2014; Schirmer & Adolphs, 2017; Schweinberger, Kawahara, Simpson, Skuk, & Zäske, 2014; Yi, Leonard, & Chang, 2019; Young & Bruce, 2011; Young, Frühholz, & Schweinberger, 2020; Yovel & Belin, 2013)

**Notes: I would probably tend to select the following eight theory papers for your paper (order chronological).** Note: With the exception of Belin et al. (2004), I have ignored pre-2010 work (in which case Ellis, Jones, and Mosdell, 1997, BJP, and Neuner & Schweinberger, 2000, Brain & Cognition, would also have been candidates)

1. Belin et al., 2004, TICS: The first layout of Belin´s “auditory face” model. Almost fully modelled to parallel Bruce & Young´s (1986) model for faces, and for that reason if not others, well connected to person perception theory
2. Belin et al., 2011, BJP: Well-cited, model with focus on both function and neuroanatomy and the “temporal voice areas” TVA
3. Kreiman & Sidtis, 2011: This is a book, and also the theory “fox and hedgehog model” is rather weak in its formulation – but the authors still get some basics right, and after all still this is from Diana Van Lancker Sidtis – one of the pioneers of voice research in the 1980s which represented precursors to the first functional models, so I feel it would be good to refer to it – especially because the 2011 book also seems to have been the very first substantial book on voice perception.
4. Young & Bruce, 2011, BJP: A 25-year overhaul of the super-influential grandmother cognitive model of Bruce & Young 1986; this paper makes more explicit consideration to voices (though admittedly not too much – a borderline case)
5. Yovel & Belin, 2013, TICS: A review focusing on coding similarities between voices and faces, with a cognitive model in the tradition of Bruce & Young 1986
6. Schweinberger et al., 2014, WIREs: Another review which takes some more explicit focus on multimodal face-voice integration
7. Young et al., 2020, TICS: A review defining the state-of-knowledge regarding merits and limitations of Belin´s auditory face model. Focuses on emotions, identity and speech, and also considers face-voice integration.
8. Lavan & McGettigan, 2024, Communications Psychology: likely very decent work to reflect the current state-of-the-art, with a few limitations especially regarding perspective on personality, too early to be much cited, and journal does not (yet) have an impact factor.

I think the above would be my top 8, but here are others you might want to consider for a reserve list, depending on the topical focus (e.g., functional vs. neuroanatomical, intact vs impaired system). If I´d want to add two more I would probably go for #1 and #8 below – but obviously judge for yourself:

1. Gainotti (2011), Neuropsychologia: Is also close to “our” Bruce-Young-Belin type of functional model, but uses the case of selective neuropsychological impairments (e.g., after stroke, focal lesion, or fronto-temporal dementia) to inform functional models of voice/person perception
2. Moerel et al. (2014), Front Neurosc: Neuroanatomy: A review on evidence from high-res (7T) fMRI research to map out functional properties of auditory primary and association cortex. Well cited – probably too specific but perhaps of interest.
3. Deen et al. (2015), Cer Cortex: Bit similar to Moerel – though original work rather than a review – and also focusing less on perceptual and more on social perception, and the superior temporal sulcus area specifically. Well cited.
4. Frühholz et al. (2016), NBR: Discusses the neural network of voice processing; is limited to emotion/affect processing, but well-cited here.
5. Schirmer & Adolphs (2017) TICS: Similar to above. Convergence and divergence of signals from face, voice, and touch. Discusses sensory and amodal representations. Review is limited to emotion perception, but well-cited here.
6. Hall et al. (2019), Ann Rev Psychol: Important review on nonverbal communication of which voices will be part – scholarly written paper in THE outlet for reviews in psychology, but perhaps not full spot on your topic
7. Yi et al. (2019), Neuron: Develops a neural population coding theory of speech, also limited to the superior temporal gyrus. Haven´t yet read this one in detail, and it might not be particularly relevant in its focus largely on speech rather than voice. But may be worth a look – and also was published in an absolute top journal.
8. Frühholz & Schweinberger (2021), Prog. Neurobiol. Probably the most extensive recent review on neural networks underlying both voice perception and voice production, with a cross-species perspective, and a focus more on neural than functional systems.
9. Bhaya-Grossman et al. (2022), Ann Rev Psychol: Important review on speech perception in superior temporal gyrus. Haven´t read this in detail. Like Yi et al (2019), this might not be very relevant due to its focus on speech rather than voice. Still, keep on the radar – I am convinced we can´t well understand voice perception while ignoring scientific progress in the neural processing of speech, as these have various inter-dependencies.

References

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding Voice Perception. *British Journal of Psychology, 102*, 711-725. Retrieved from WOS:000296085900003

Belin, P., Fecteau, S., & Bedard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences, 8*(3), 129-135. Retrieved from ISI:000220326900009

<http://ac.els-cdn.com/S1364661304000257/1-s2.0-S1364661304000257-main.pdf?_tid=edf5252e-a871-11e4-b753-00000aab0f26&acdnat=1422616980_aa42cb2fcda3623cc07fac1a55f05a3e>

Bhaya-Grossman, I., & Chang, E. F. (2022). Speech Computations of the Human Superior Temporal Gyrus. *Annual Review of Psychology, 73*, 79-102. doi:10.1146/annurev-psych-022321-035256

Deen, B., Koldewyn, K., Kanwisher, N., & Saxe, R. (2015). Functional Organization of Social Perception and Cognition in the Superior Temporal Sulcus. *Cerebral Cortex, 25*(11), 4596-4609. doi:10.1093/cercor/bhv111

Frühholz, S., & Schweinberger, S. R. (2021). Nonverbal auditory communication - Evidence for integrated neural systems for voice signal production and perception. *Progress in Neurobiology, 199*, 101948. doi:10.1016/j.pneurobio.2020.101948

Fruhholz, S., Trost, W., & Kotz, S. A. (2016). The sound of emotions-Towards a unifying neural network perspective of affective sound processing. *Neuroscience and Biobehavioral Reviews, 68*, 96-110. doi:10.1016/j.neubiorev.2016.05.002

Gainotti, G. (2011). What the study of voice recognition in normal subjects and brain-damaged patients tells us about models of familiar people recognition. *Neuropsychologia, 49*(9), 2273-2282. Retrieved from WOS:000293611600001

Hall, J. A., Horgan, T. G., & Murphy, N. A. (2019). Nonverbal Communication. In S. T. Fiske (Ed.), *Annual Review of Psychology, Vol 70* (Vol. 70, pp. 271-294).

Kreiman, J., & Sidtis, D. (2011). *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception*. Chichester: Wiley-Blackwell.

Lavan, N., & McGettigan, C. (2023). A model for person perception from familiar and unfamiliar voices. *Communications Psychology, 1*(1). doi:<https://doi.org/10.1038/s44271-023-00001-4>

Moerel, M., De Martino, F., & Formisano, E. (2014). An anatomical and functional topography of human auditory cortical areas. *Frontiers in neuroscience, 8*. doi:10.3389/fnins.2014.00225

Schirmer, A., & Adolphs, R. (2017). Emotion Perception from Face, Voice, and Touch: Comparisons and Convergence. *Trends in Cognitive Sciences, 21*(3), 216-228. doi:10.1016/j.tics.2017.01.001

Schweinberger, S. R., Kawahara, H., Simpson, A. P., Skuk, V. G., & Zäske, R. (2014). Speaker perception. *Wiley Interdisciplinary Reviews-Cognitive Science, 5*(1), 15-25. Retrieved from WOS:000328560300002

Yi, H. G., Leonard, M. K., & Chang, E. F. (2019). The Encoding of Speech Sounds in the Superior Temporal Gyrus. *Neuron, 102*(6), 1096-1110. doi:10.1016/j.neuron.2019.04.023

Young, A. W., & Bruce, V. (2011). Understanding person perception. *British Journal of Psychology, 102*, 959-974. Retrieved from WOS:000296085900018

Young, A. W., Frühholz, S., & Schweinberger, S. R. (2020). Face and voice perception: Understanding commonalities and differences. *Trends in Cognitive Sciences, 24*(5), 398-410. doi:10.1016/j.tics.2020.02.001

Yovel, G., & Belin, P. (2013). A unified coding strategy for processing faces and voices. *Trends in Cognitive Sciences, 17*(6), 263-271. doi:10.1016/j.tics.2013.04.004