

Musician advantage for speech-on-speech perception

Deniz Başkent^{1,2}, Michael Chesnay¹, Nikki Tahapary¹, and Etienne Gaudrain^{1,2,3}

¹Department of Otorhinolaryngology / Head and Neck Surgery, University Medical Center Groningen, The Netherlands

² Research School of behavioral and Cognitive Neurosciences, University of Groningen, The Netherlands

³ Lyon Neuroscience Research Center, Auditory Cognition and Psychoacoustics Team, CNRS UMR 5292, Lyon, France

E-mail: d.baskent@umcg.nl

Introduction

Musicians show advantages over non-musicians in a number of auditory perceptual tasks. Such advantage may be due to better processing of acoustic features, such as fundamental frequency (F0), which is also a primary dimension in music, or due to enhanced cognitive abilities, such as better attention abilities or extended working memory capacity, at least in the auditory modality.

Evidence for a transfer of musician advantage to better perception of speech in noise has been mixed (Fuller et al, 2014a; Parbery-Clark et al, 2009; Ruggles et al, 2014). These studies have primarily used noise as background masker, which may be a relatively less efficient condition to identify potential musician advantage, as perception of speech in noise does not seem to rely on fine F0 processing. Further, this task is more heavily dominated by energetic masking and less so by informational masking, hence, cognitive abilities likely play a minimal role.

Speech on speech perception, on the other hand, has been shown to directly depend on F0 differences between the two competing voices (e.g. Darwin et al, 2003) and involves informational rather than energetic masking, which mobilizes more cognitive resources (e.g. Zekveld et al, 2013). Therefore, we hypothesized that speech-on-speech could be a more suitable test condition than speech-in-noise to investigate the potential musician advantage for speech perception.

Research Questions

Is there a musician advantage for speech on speech perception?

If yes, does it mainly come from better processing of F0?

Methods

- Musician definition: To have started musical training before/at the age of 7, have had at least 10 years of musical training and have been active with music the last 3 years before testing.
- Non-musician definition: Not meeting the inclusion criteria for musicians, not having received musical training within 7 years before the study.
- 18 musicians (19-25 yrs) and 20 non-musicians (19-27 yrs)
- Identification of sentences presented concurrently with a masker sentence. Target to masker ratio (TMR)=- 6 dB.
- Voice separation by shifting masker sentences (MATLAB STRAIGHT TANDEM) in two dimensions of voice characteristics:
 - Fundamental frequency ($\Delta F0$): 0, 4, 8 st
 - Vocal tract length (ΔVTL): 0, 0.75, 1.5 st

References

- Darwin, C. J., Brungart, D. S., & Simpson, B. D. (2003). Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. *JASA* 114(5), 2913–2922. doi:10.1121/1.1616924
- Fuller, C. D., Galvin, J. J., Maat, B., Free, R. H., & Başkent, D. (2014a). The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations? *Fr Neuroscience*, 8, 179. doi:10.3389/fnins.2014.00179
- Fuller, C. D., Galvin, J. J., Free, R. H., & Başkent, D. (2014b). Musician effect in cochlear implant simulated gender categorization. *JASA* 135(3), EL159–165. doi:10.1121/1.4865263
- Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing*, 30(6), 653–661. doi:10.1097/AUD.0b013e3181b412e9
- Ruggles, D. R., Freyman, R. L., & Oxenham, A. J. (2014). Influence of musical training on understanding voiced and whispered speech in noise. *PLoS One*, 9(1), e86980. doi:10.1371/journal.pone.0086980
- Zekveld, A. A., Rudner, M., Johnsrude, I. S., & Rönnerberg, J. (2013). The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. *JASA* 134(3), 2225–2234. doi:10.1121/1.4817926

Funding

Rosalind Franklin Fellowship from the University Medical Center Groningen, VIDI grant 016.096.397 from the Netherlands Organization for Scientific Research (NWO).

Results

Fig 1: Percent correct scores shown for fixed ΔVTL in each panel, as a function of $\Delta F0$.

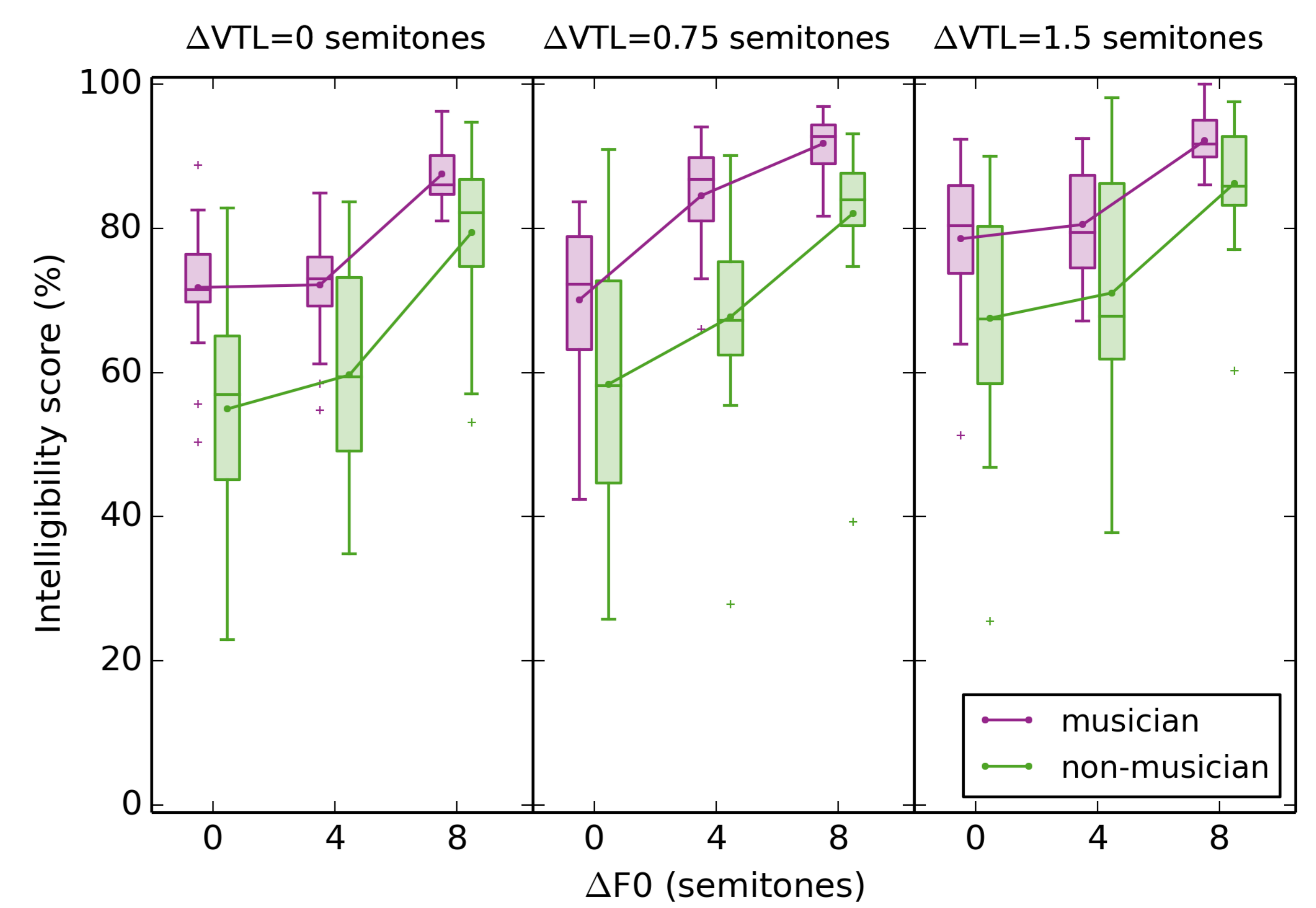
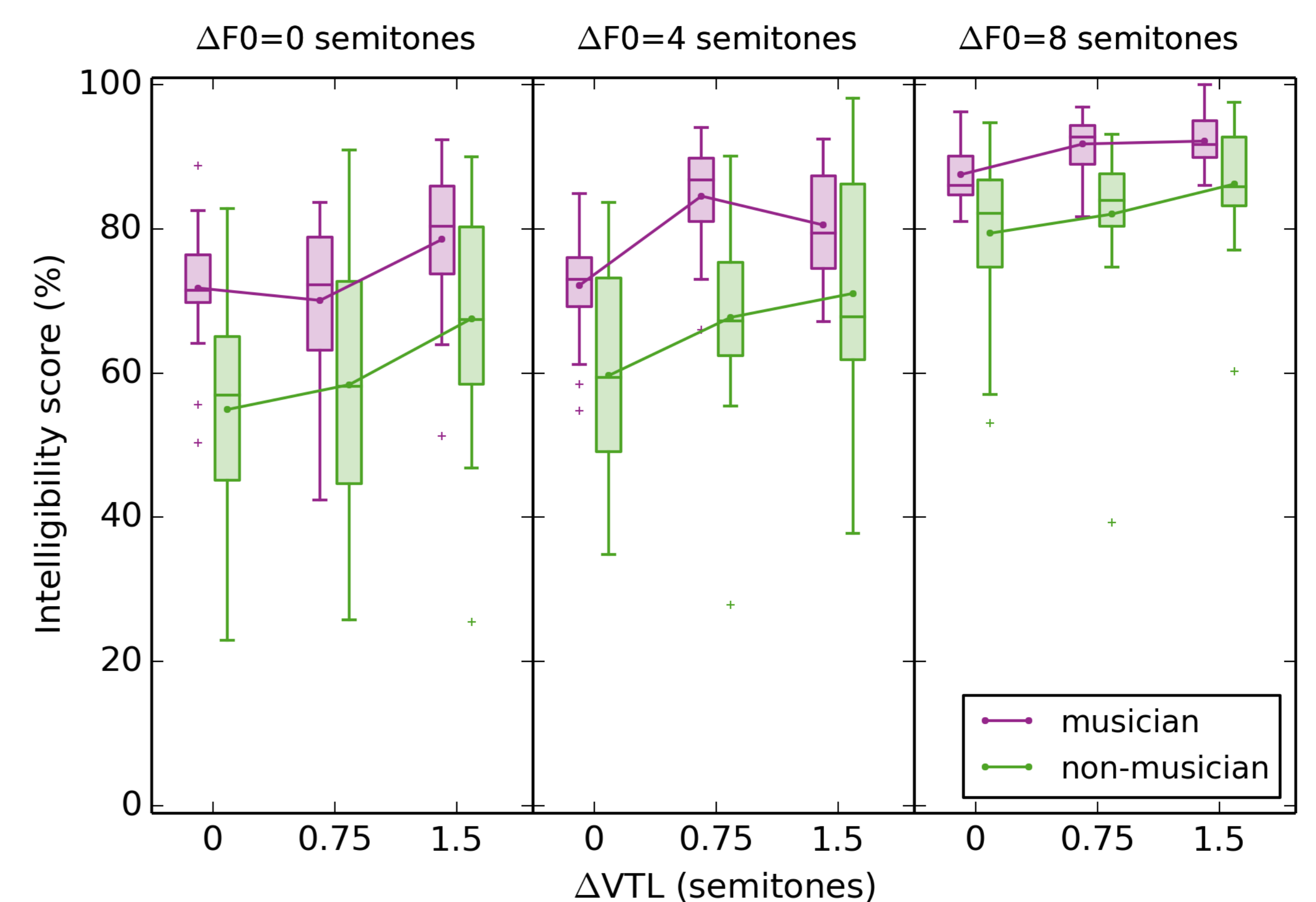


Fig 2: Percent correct scores shown for fixed $\Delta F0$ in each panel, as a function of ΔVTL .



A generalized linear mixed model with subject as random intercept showed that both VTL and F0 had significant effects on intelligibility. Moreover, the musicians performed better than the non-musicians [$z=-4.07$, $p<0.001$]. However, the musicians were not better than the non-musicians at using voice characteristics.

Conclusions

- As observed before (Darwin et al., 2003), F0 and VTL differences between the competing sentences both contributed to improve intelligibility.
- The musicians unambiguously performed better than the non-musicians.
- However this difference in performance was independent of the voice cue difference: the two groups benefited similarly from F0 and VTL differences. This finding is against the previous observation of Fuller et al (2004b) that musicians utilized ΔVTL less and $\Delta F0$ more efficiently than non-musicians, but in a different task, namely talker gender identification.
- Speech on speech perception relies only partially on separation of voices, while the rest of meaning extraction is achieved via the means of cognitive processing. Further, it relies more on information masking than other forms of non-speech maskers, again indicating more involvement of cognitive skills. Hence, our results perhaps reflect the better auditory cognitive skills in musicians.

