Supplementary Materials for "L2 difficulties in the perception of Mandarin tones: Phonological universals or domain-general aptitude?"

Chao Zhou[1] & João Veríssimo[1]

[1] Center of Linguistics, School of Arts and Humanities, University of Lisbon

Paper in press at *Bilingualism: Language and Cognition*

This work has been funded by the Fundação para a Ciência e a Tecnologia (FCT, Foundation for Science and Technology), grant UID/00214: Center of Linguistics of the University of Lisbon. We thank Ling Li for help in recording the stimuli, and the audiences of the Architectures and Mechanisms for Language Processing (AMLaP) 2023, Manchester Phonology Meeting (mfm) 2023, and Encontro Nacional da Associação Portuguesa de Linguística (ENAPL) 2022 conferences for helpful discussions.

Correspondence concerning this article should be addressed to João Veríssimo, Faculdade de Letras da Universidade de Lisboa, Alameda da Universidade, 1600-214 Lisboa, Portugal. E-mail: jlverissimo@edu.ulisboa.pt

Appendix S1. Prior sensitivity analyses

Bayes factor analyses can be heavily influenced by the choice of prior distribution on the parameters of interest, so it is considered good practice to investigate claims about particular hypotheses under a range of reasonable priors (Depaoli et al., 2020; Schad et al., 2021; Sinharay & Stern, 2002). To ensure that our conclusions were not determined by the use of a weakly informative Normal(0, 2.5) prior on the fixed effects, the most important effects reported in this paper were tested in a sensitivity analysis that made use of four other priors. More specifically, the comparison between identical and non-identical pairs (which diagnosed the OCP) and the comparisons between tone pairs (which diagnosed the TMS) were additionally tested under two narrower priors (Normal(0, 1.25) and Normal(0, 0.625)), whose SDs were one-half and one-fourth of the SD of the default prior, and two wider priors (Normal(0, 5) and Normal(0, 10)), with twice and four times the SD of the default prior. The smaller the prior SDs, the less extreme the differences between conditions were assumed to be, so that altogether, these distributions capture a wide range of a priori beliefs about the possible magnitudes of effects.

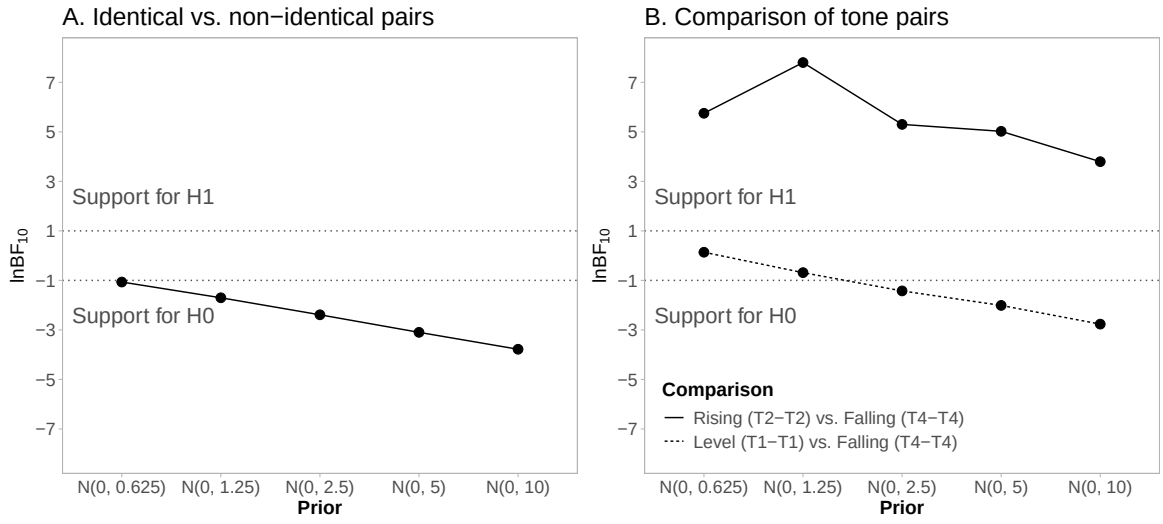

Figure S1. Prior sensitivity analyses for the comparison between accuracies in identical and non-identical pairs (left panel) and accuracies in different tone pairs (rising vs. falling and level vs. falling), under five different Normal priors. Values of $\ln BF_{10}$ greater than 1 support the existence of an effect (i.e., H1) and values smaller than -1 support its absence (i.e., H0) (Kass & Raftery, 1995).

Figure S1A shows the logged Bayes factors in favour of the alternative hypothesis ($\ln BF_{10}$, Kass & Raftery, 1995), under the five priors discussed above (including the default prior reported in the main paper), for the comparison between identical and non-identical pairs. The results show that the null hypothesis received support under all five priors (values of $\ln BF_{10}$ greater than 1 support the existence of an effect and values smaller than -1 support its absence). Thus, we have obtained evidence for equal accuracy of tonal identification in identical and non-identical pairs, against the OCP.

Figure S1B shows the results of the same sensitivity analysis for the comparisons between different tonal pairs (rising vs. falling and level vs. falling). The results show that there was evidence for a difference between rising (T2-T2) and falling tones (T4-T4), irrespective of the prior distribution that was employed, with rising tones being much less accurately identified than falling tones (see main paper). The conclusions that can be drawn for the comparison between level (T1-T1) and falling tones (T4-T4) depend on the width of the prior distribution that was employed: for the default Normal(0, 2.5) prior and the two wider priors, the Bayes factors supported equal identification accuracies, whereas with the two narrower priors, the results are essentially inconclusive, with no support for equality nor for a difference between level and falling tones. Thus, the TMS was only partially supported in our data, with participants being much less able to correctly identify rising tones, but performing similarly for falling and level tones.

Appendix S2. Analysis of constituent tones

An anonymous reviewer pointed out another plausible account for the lack of OCP effects in L2 perception. In our OCP test, contour tones were examined as holistic units (Yip, 2002; Zhang, 2016), according to which the OCP was expected to operate over tone pairs like T2-T2 and T4-T4. Nevertheless, L2 learners might represent contour tones as compositions of level tones (e.g., T4 as H and L), according to which the OCP would militate against adjacent identical constituent tones. Zhang (2016) tested this hypothesis in her production studies, but did not find evidence for it.

We performed a new statistical analysis to examine the influence of OCP on constituent tones, by fitting an analogous OCP model, but in which pairs in the OCP-violating condition were T1-T1, T1-T4, T2-T1, T2-T4, and T4-T2. The results showed that participants were actually better at identifying tone sequences in which the constituent tones violated the OCP than at identifying those that did not (b = `0.66 [0.35, 0.98]`, $\ln BF_{10}$ = `5.7`). Consistent with Zhang's (2016) results, we did not obtain evidence for the hypothesis that the OCP applies to constituent tones in L2 learners of Mandarin.

References

Depaoli, S., Winter, S. D., & Visser, M. (2020). The importance of prior sensitivity analysis in Bayesian statistics: Demonstrations using an interactive Shiny App. *Frontiers in Psychology*, *11*. https://doi.org/10/gmptn8

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10/gdnbw3

Schad, D. J., Betancourt, M., & Vasishth, S. (2021). Toward a principled Bayesian workflow in cognitive science. *Psychological Methods*, *26*(1), 103–126. https://doi.org/10/ghbtt6

Sinharay, S., & Stern, H. S. (2002). On the sensitivity of Bayes factors to the prior distributions. *The American Statistician*, *56*(3), 196–201. https://doi.org/10/fdrhk8

Yip, M. (2002). *Tone*. Cambridge University Press.

Zhang, H. (2016). Dissimilation in the second language acquisition of Mandarin Chinese tones. *Second Language Research*, *32*(3), 427–451. https://doi.org/10/f8w822