

# Response to reviews

28/11/2018

You mention on line 496 that “language does not seem to be an informative parameter”. As I read your paper, you analyzed speakers from both languages together, but you should probably say so in E. Data processing and statistical analysis.

*I've added "Speakers of both Italian and Polish were analysed together (models were fitted to a single dataset with data from both languages)." on lines XXX.*

Line 28 It is not clear what “C2” refers to.

*Replaced "C2" with "closure".*

Line 127 “The duration of the release to release 2interval is stable across voicing contexts.” It is not clear to me why this interval is interesting and relevant. The word “stable” probably means “non-significant differences”, but this can simply be due to a lot of variability.

*I have rephrased that sentence and added text: "Finally, the temporal distance between two consecutive stop releases in CVCV words is not affected by the voicing of the second consonant. According to a Bayes factor analysis (Coretta 2019), the duration of the release to release interval is not affected by the voicing of C2. Within this interval, the timing of VC boundary (the vowel offset/onset of stop closure) produces differences in the respective durations of vowel and closure, following a mechanism of temporal compensation ([citations]). A later closure onset results in a long vowel and a short closure, while an earlier closure onset corresponds to a short vowel and a long closure. Since the closure of voiceless stops is longer than that of voiced stops, it follows that vowels are shorter when followed by the former than when followed by the latter. As more thoroughly discussed in Coretta (2019), the release to release interval per se does not have a special status, but rather it has been used as a proxy to acoustic temporal stability more in general. The aspects of the compensatory account proposed in Coretta (2019) that are relevant to the current study are that (1) C2 voicing does not affect the interval that includes V1 and C2, and that (2) the placement of the VC boundary determines the duration of both V1 and C2 closure. Future studies are warranted to investigate the production and/or perceptual reasons behind both (1) and (2)."*

Line 380 “...that the onset of the forward gesture of the root is timed not relative to the stop closure, but rather relative to a fixed time point preceding the closure.” I would be very careful to talk about “fixed time points” in speech. You don't elaborate. Since the tongue is always moving during speech, movement onsets can be hard to locate, and you don't do so. You make a similar statement on line 443.

*I've rephrased: "but rather relative to an acoustic/articulatory event preceding the closure (what this event might be should be further investigated)." and "In order to obtain such correlation, the articulatory onset of the advancement gesture (during the articulation of the vowel) should be at a stable distance from an earlier reference point (like the vowel onset or the preceding consonant offset) such that the timing of consonant closure will create the correlation seen in the data. Although ideally the timing of the onset of the advancing gesture relative to a preceding articulatory landmark should not be affected by the voicing of C2, the velocity of the gesture itself could be different depending on the voicing of the following consonant."*

Line 448 Here, you discuss the relationship between movement amplitude and velocity. There is a strong correlation between them for both speech and non-speech movements.

Footnotes #2, 3, 5, and 6 can be incorporated into the text.

*They have been incorporated in the text body.*

## References

Lampp & Reklis, 2004, appears to be an abstract, thus not very useful, and should be removed.

Ohala, 2011, is available here> <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/SpecialSession/Session1/Ohala/Ohala.pdf>

Strycharczuk & Scobbie, 2015, is available here> <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0309.pdf>

Sprouse et al, 2008, is available here> <https://issp2008.loria.fr/Proceedings/PDF/issp2008-101.pdf>

*I have fixed the references.*

General comments - This paper has improved very much in clarity. The author has confirmed that measures of tongue root position were reversed, and has made efforts to address my concerns regarding measurement method. I find I still have the same reservations relating to ability to take accurate, fine-grained measures using the UTI method - there are issues relating to the fuzziness of the UTI-imaged tongue surface and the accuracy of the automatic spline fitting with manual correction. I can't see any other way around this than for the author to acknowledge these potential problems either in the method or discussion section, e.g. in B. Estimates of tongue root displacement, where the author discusses the small difference in adjusted mean root displacement (0.77-1.0mm) between the two following consonant conditions. I therefore suggest revise and resubmit.

#### Detailed comments

1. L27-30 "Furthermore, the results of this study indicate that a comparatively later C2 onset for voiced consonants results in a longer preceding vowel duration which, in turn, results in greater tongue root advancement at C2 onset." It seems obvious that a later C2 onset for voiced consonants results in a longer preceding vowel duration. Maybe you need to rephrase as "... indicate that a comparatively later C2 onset for voiced consonants, resulting in a longer preceding vowel duration, correlates with greater tongue-root advancement at C2 onset."

*Yes, thank you. I rephrased it as suggested.*

2. L94 "exclusive to voiced stops"

*Corrected.*

3. L94-96 "While this gesture is not exclusive of voiced stops and it is sometimes implemented even in the absence of vocal fold vibration, tongue root advancement seems to be a robust correlate of (phonological) voicing." Given what you say in the first two lines, the word "robust", does not seem right, maybe "tongue root advancement is strongly associated with tongue-root advancement..."

*Yes, it makes sense. I rephrased as suggested.*

4. L124 "Independent of language, some speakers have a greater effect (of following consonant voicing on vowel duration) and others a small or negligible effect".

*Rephrased.*

5. L146-7 4-line gap in manuscript.

*That is just LaTeX being too strict with widow lines. This will fix itself in the final manuscript.*

6. L223-5 How did you deal with any aspiration occurring after the initial plosive consonant. Did it occur? Was it included in the vowel duration calculation?

*Some voiceless-post aspiration occurs in Polish C1 stops (mean = 35 ms, SD = 28.5). Post-aspiration duration nor burst duration were included in the measurement of vowel duration as defined in this study. I have added this sentence: "The burst and any eventual voiceless post-apiration of C1 are not included in the duration of the V1."*

7. L246 "oral track" → "vocal tract"

*Fixed.*

8. L253-5 "To facilitate interpretation of the displacement values, the sign of these was flipped so that higher values indicate a more advanced tongue root (greater tongue root advancement) after Kirkham and Nance (2017)."

*Added.*

9. Figure 2 caption “-after z-scoring normalization, the sign is flipped so that greater values indicate greater tongue root advancement).”

*Added.*

10. L284 “offest” → “offset”

*Fixed.*

11. L303, 305, 322, 323 change “,” to “;”

*Fixed.*

12. L316-7 “However, the magnitude of the movement is greater in the former, and begins earlier in the vowel”??

*Since the duration of vowels is normalised, we cannot say that the tongue root advancement gesture begins earlier in vowels followed by voiced stops than in vowels followed by voiceless stops.*

13. L320 V1 offset/C1 onset → C2 onset

*Fixed.*

14. Fig. 5 Tongue root position is labelled in (ms), should be (mm).

*Fixed.*

15. Fig. 6 is in an odd location. i.e. in the Discussion section, rather than the results.

*Figures in the manuscripts are floats so they are placed in the best position possible by LaTeX. I forced the placement of this specific figure in the manuscript to be the same as that in the TeX source file.*

16. Discussion §IV.A. I think there has to be an acknowledgement here that there is a small adjusted mean (0.77mm-1.0mm) difference in ATR at V1 offset between the voiced and voiceless C2 contexts. Also some statement has to be made about the potential errors involved in spline fitting, especially where manual correction is involved.

*I have rephrased as: “Unsurprisingly, the position of the tongue root at vowel offset is 0.77 mm (SE = 0.35) more front when the following stop is voiced than when the following stop is voiceless in both surveyed languages (see Section ... for a discussion about the magnitude of the difference and potential errors related to spline fitting).” And in Lines ... “A note of caution is due, since the actual error rate of the automatic tracker used for spline fitting is not known, and manual correction might have affected the splines (although a relatively small number of tokens had to be manually corrected).”*

17. L379-80 “could be interpreted as to indicate that...” → “could indicate that”

*Fixed.*

18. P34 Why is Table 1 here instead of on P10?

*Moved to Materials section.*

19. L358 “such a small degree of advancement in voiceless lingual stops” According to your findings, the degree of advancement before voiceless lingual stops is only marginally less than in voiced stops.

*Yes. I’ve removed “such a small degree of”.*

20. L379-388 I found the argument of this paragraph very hard to follow. Perhaps some diagrams would be helpful.

*I’ve added Figure 7.*

21. L560 “See ??”

*Fixed.*

22. L579 Page break missing.

*I have tried hard to add a page break before the bibliography by changing the TeX source but it won't work, so I trust this will be fixed in the final manuscript.*