*Revisions for "Using optical flow analysis on ultrasound of the tongue to examine phonological relationships" by K.C. Hall, H.Smith, K. McMullin, B. Allen and N. Yamane*

1. Both reviewers comment on the fact that the segmental contexts of the stimuli were not controlled, and that surrounding context might be the source of the apparent differences in tongue movement in closed vs. open syllables. We have followed their suggestions of examining subsets of the data that are more closely matched in phonetic context, and especially to use subsets that involve contexts that themselves involve as little tongue movement as possible, i.e., labial contexts. Doing so reveals that the exact same pattern of results holds, reducing the concern of context being the driving force here; we appreciate this suggestion, because it certainly enhances our argument. We have included results from two subsets in the revised article for clarity.

2. Both reviewers ask for clarification on the actual calculation of the magnitude measure, with reviewer E being especially interested in a comparison of the current technique to other OFA implementations. We have provided the specific equation (from the Barbosa et al. 2008 paper originally cited) and given discussion for why this particular strategy of dimensionality reduction is appropriate for the specific research question being addressed here.

3. Reviewer C says: "All of the frames that were identified as being part of a target vowel articulation were collected from the data file and subjected to a z-score transformation. Thus, the dependent measure appears to be the relative amount of pixel movement observed from one frame to the next for each speaker. These values were then grouped by vowel and averaged. I do not know whether or not that is an appropriate measure to use, but I would like to see a justification from the authors for why that is preferable over a number of other possibilities." We have provided explicit justification for each individual step of this process in the article itself. The reviewer suggests an alternative approach, in which the magnitudes from each frame are summed rather than averaged. We don't explicitly address this possibility in the article, but the reason that this is not a viable alternative is that there are differing numbers of frames per vowel, such that longer vowels would simply appear to have more movement if this approach were adopted (and this latter point is discussed explicitly as part of the discussion of the chosen technique).

4. Reviewer C: "It is also not clear to me what happens when there is a movement in opposite directions within a frame and across an analysis window." We have explicitly addressed this question; the measures are that of magnitude, and do not include direction, so movement is additive even when it is in the opposite direction.

5. Reviewer C: "It may be useful to list the number of words compared for each vowel in Figure 2." We have not listed the specific number for each plot, but have explicitly given the information (as surmised by the reviewer) about how many measurements are included.

6. Reviewer C asks for more detail about the audio recording. We are not sure exactly which information is being asked for, but we have tried to be as detailed as possible.

7. Reviewer C points out that the level of precision for determining vowel boundaries in Praat is higher than is actually possible for finding the corresponding frames in the video; we acknowledge this.

8. Reviewer E correctly points out that the participants in the study likely were speakers of a dialect where there is an [ɔ] / [ɑ] merger. While this isn't a direct problem because the vowels of interest are the tense vowels, i.e., [o], it does mean that the comparison set for calculating the degree of contrast should include all [ɑ] vowels, i.e., we should consider the possibility that the relevant contrast is [o] / [ɑ], not [o] / [ɔ]. The calculations have been re-done to include the possibility of a merger, and both scenarios are explicitly discussed in the article. As with taking a subset of the original data, this change does not materially affect the results or the argumentation.

9. Reviewer E also asks for a clarification of the role of duration, which we have provided. The measurement we provide is the average magnitude per frame, so duration is not a factor.

10. Reviewer E's smaller comments:
    a. A reference for differing calcification rates has been added.
    b. The details of the chair are largely irrelevant and have been omitted.
    c. We agree that 18" was too distant for the microphone; this was an unfortunate error on our part. This is the reason that no acoustic analysis was ever provided in the article.
    d. We have clarified that yes, it is the entire ultrasound image that was included.
    e. We have included more detailed information about the placement of the probe.