



School of Psychology
Science Centre Level 2, Building 302,
23 Symonds Street Auckland, New Zealand
(+64) 21-672-013

30 April, 2025

RE: *A Programmatic Stage 1 Registered Report of global song-speech relationships replicating and extending Ozaki et al. (2024) and Savage et al. (2025)*
(<https://doi.org/10.31234/osf.io/c2dba>)

Dear Dr. Logan,

We are delighted that the reviewer (Dr Moscoso del Prado Martin) was satisfied with our previous revision and appreciate your invitation to revise and resubmit our Programmatic Stage 1 Registered Report protocol to address your additional comments. You will see that we have tried to implement most of your suggested changes, but we respectfully prefer to keep the broader Programmatic context in the Introduction and Methods section in some cases to provide enough context, at least for this Stage 1 protocol. We have appended a version with tracked changes to this response letter for your convenience.

We feel that the review process has substantially improved our Stage 1 protocol. We hope you will find the revised protocol ready for In-Principle Acceptance and Stage 2 data collection.

Sincerely,

Patrick Savage (on behalf of the authors)

Full reviews/decision:

Review by Fermin Moscoso del Prado Martin, 09 Apr 2025 15:49

I think this is an interesting and well-designed study. The authors have satisfactorily address my suggestions in the previous round.

We appreciate your enthusiasm!

Recommender's decision/summary (Corina Logan):

Decision for round #2 : *Revision needed*

Revise

Dear Dr Savage and co-authors,

Thank you for your revised submission and for responding to feedback. The methods section is greatly improved in terms of having more information about this research, rather than having to read methods sections from two other articles to get the full picture. And the introduction is more fleshed out in terms of what is driving this research and what it means.

Thanks for your guidance in achieving this!

Now that these details are included in the manuscript, it is easier to assess the proposed work.

The main issue now is that the Stage 1 reads like an outline that is addressing points in the author guidelines at PCI RR, rather than being a full article that follows the author guidelines without explicitly discussing the guidelines. In its current state, it is unlikely that authors will be able to copy and paste the intro and methods into their Stage 2s without major revisions to the text, which is not permitted at Stage 2 (see section 2.10 in the PCI RR author guidelines https://rr.peercommunityin.org/help/guide_for_authors#h_97949820420921613309536944). For example, I appreciate that you made the section “Equitable coauthorship in global collaboration”. However, now that I see it in the context of this draft, I think it should be put into supplementary material because it is a commentary on your workflow and how you will divide up the labor that goes into the Stage 2 articles - the process behind the articles. This is useful for people to know because you are developing a new way of working together, but this is not the focus of your research and I think it distracts from the research in the main text.

We experimented with moving this section to supplementary material as suggested, but found that it resulted in burying some key aspects of the motivation for the study. With great respect for your suggestion (appreciating that you are one of the only people in the world who has actually started implementing a Programmatic Registered Report), we prefer to leave this sub-section in the Introduction. We feel that this equitable collaboration model is in fact one of the key motivations for this series of studies, and the rationale for each individual study is hard to fully appreciate without the context of the full collaboration. We also note that, while it is true that major deviations in hypotheses, etc., are not permitted at Stage 2, PCI-RR's guidelines for Programmatic RRs (section 2.15) do state that *"Because each Stage 2 submission relates only to a portion of the approved Stage 1 protocol, it is understandable that authors may need to make greater structural amendments to the introduction and method sections of the Stage 2 manuscripts compared with a regular RR (including only the relevant content)."*

Thus, if you feel strongly that this section (and/or other sections referring to the broader Programmatic plan) does not belong in the resulting Stage 2 reports, it should be acceptable to either move these to Supplementary Material and/or delete them during Stage 2, but we would prefer to keep them at least in this Stage 1 protocol.

Other elements that make the Stage 1 look more like an outline than a final draft are along the following lines... Referring to PCI RR's programmatic track in the abstract and intro should be removed, and instead discuss how the outputs will be conducted and how this will benefit the research.

We have removed explicit references to PCI-RR, while keeping reference to Programmatic Registered Reports, since this is a crucial part of the design but is not yet widely known by most readers.

Instructions to recommenders and reviewers can be put in author responses and cover letters, or you might want to put this in supplementary material because you are developing a new way of working together that others might want to use.

We have edited the protocol text to strike a better balance here (only quoting a couple of key phrases when needed, such as that unregistered exploratory analyses must be "justified, methodologically sound, and informative", otherwise just citing the guidelines without explicitly quoting them).

Table 1 should be included in the main text (maybe in the methods?) rather than as a separate page that has no section title between the abstract and introduction (also landscape orientation for this table would make it fit better on the page).

We have moved this table to between the Introduction and Methods (now Table 2, in landscape).

The “Hypothesis” section should be incorporated into the regular text without the word “Hypothesis” denoting it as a separate section.

We experimented with this, but found that if we are keeping the “Equitable coauthorship in global collaboration” sub-section this hypothesis sub-section still requires it’s own sub-heading or the structure becomes confusing.

The Hypothesis section has methods in it that should be moved to the methods section.

We moved the following two paragraphs from the Introduction to the Methods section:

Except for the Stage 2 output combining all studies (#27 in Table 2), each Stage 2 will focus its confirmatory analyses on the results of its own analysis of its own focus language. #27 will replicate Ozaki et al.'s (2024) cross-linguistic meta-analysis approach to analyse average trends across all languages, which can be compared with the results of each individual Stage 2 reports #1-26 to achieve a much broader evaluation of the cross-linguistic replicability and generalisability of Ozaki et al.'s (2024) original results. Comparison of specific differences between languages will be reserved for exploratory analysis (since statistical power for such comparisons will be limited by the relatively small sample size of n=15-30 participants per language).

To ensure maximal consistency across Stage 2 reports, all Stage 2 reports will restrict their confirmatory analyses and statistical hypothesis testing to only these three hypotheses. They are welcome and encouraged to explore additional unregistered exploratory analyses, but these must be “justified, methodologically sound, and informative”³⁹.

Explicit references to things like PCI RR’s criterion 2D should be removed because the content in the intro and methods section should state what will be done (which will be in accordance with this criterion, but without explicitly pointing it out).

Done

In the methods section, don’t refer to Stage 2s or first authors of Stage 2 articles and so on, just say what the minimum sample size is for each population/site (whatever term you want to use) and how each population will be coded (e.g., line 228).

We have implemented this as much as seemed feasible (e.g., deleting “for the 26 proposed single-site Stage 2 reports” from the former line 228 [now line 334]), while still preserving the overall Programmatic design as discussed above.

There are many other elements along these lines, including in the methods section, but hopefully this feedback gives you enough of an idea to be able to change the rest of these elements.

In other words, this Stage 1 should be about the research you are proposing to conduct and not a discussion about your workflows (you can put the latter in supplementary material). Think of this Stage 1 as being the final Stage 2 article (but without the results and discussion sections) - this will be what ends up as the final abstract and intro and methods (of course, modified for each Stage 2 article). I have several further comments; please see below.

As discussed above, we believe that each final stage 2 article will make most sense if most of the full Programmatic context is explained in the Introduction and Methods, and the Results and Discussion sections focus on the results for the relevant site(s). For example, we propose keeping the full list of sites in Table 1 and Fig. 1 in each Stage 2 report, but highlighting the relevant site on each map/table (e.g., using colours/bold).

The track changes file doesn't show individual insertions and deletions, and some changes don't show up at all. For example, it looks like the entire abstract and whole paragraphs in the intro and methods are new. However, when I compare the previous version with the current version in Draftable, I can see that there were several insertions and deletions in both. Please make sure to have clear and correct track changes files, which will help the review process go faster.

Apologies, we used Word's “compare documents” function to create the tracked changes version, which can cause issues like this. We have checked the current version and tried to make it as clear as possible (e.g., accepting its attempts to show replaced images even when the images are identical). We couldn't always find a perfect way to distinguish between old text that has been moved vs. new text, but we think this version does this much better than last time, and hope you find it clearer and more helpful.

I look forward to receiving your revised submission.

All my best,

Corina

Other comments:

Line 181 - how did you empirically determine the optimal sample size? What is the minimum sample size?

The minimum sample size was already specified below (“We will specify a minimum of 10 participants per site”). We have added the following text clarifying the optimal sample size determination:

Note that this optimal sample size was determined by combining a quantitative survey asking each potential collaborator to list the maximum number of participants they could feasibly recruit with qualitative, holistic discussions via a series of Zoom meetings with collaborators.

Figure 2 and legend needs panel numbers

Thanks for the suggestion! Here is the updated figure/legend:






A) Experiment setup (5-10 participants)	Condition	1st phrase	2nd phrase	etc.
	B) Alternating singing	 1) I don't know whoaaaa	 2) Why does love do this to me?	...
	C) Conversation	 1) I'm feeling a bit tired but enjoying the sunshine...	 2) Yeah, was great to eat lunch in the park!	...

Figure 2. Schematic overview demonstrating an example of the two conditions analysed in confirmatory analyses. Here, only the first two participants are shown singing or speaking sequentially, but the total number of participants will be between 5-10 per experiment (A). Text columns #1 and #2 represent the first and second phrases of alternating singing (B) / speaking (C) conditions. This example shows lyrics for “Why Does Love Do This To Me?”, the song chosen for participants using New Zealand English, and hypothetical conversation based on the ice-breaker prompt “How is your week going?”, but note that the actual song and conversation prompt will be different (and generally in a different language) at each site. (See Savage et al.²'s Fig. 1 for an illustration of the lyric recitation and synchronised singing conditions not included in the current confirmatory analyses.)

Line 253 and throughout the manuscript - add year to the in text citation

Done.

Line 403 - add a citation for distinguishing between thresholds

Interrater reliability - I am not sure what the standard threshold is in the fields of linguistics and music, however, for comparative cognition, which is my field, ICCs must be 0.90 or higher. I found an article on “A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research”. They suggest: “As a rule of thumb, researchers should try to obtain at least 30 heterogeneous samples and involve at least 3 raters whenever possible when conducting a reliability study. Under such conditions, we suggest that ICC values less than 0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability.”² (Koo and Li 2016)

Therefore, I recommend raising your passing threshold from 0.60 to at least 0.75, but 0.80 would be better. This is a crucial piece because it is an indicator of the quality of the data you will be analyzing, and what goes into the models needs to be as clean as possible to obtain more rigorous results from the model.

Reference: Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016 Jun;15(2):155-63. doi: 10.1016/j.jcm.2016.02.012.

Koo & Li’s guidelines are more conservative than many others (e.g., Landis & Koch, 1977; Cicchetti, 1994, who suggest that 0.60-0.74 is “good” and 0.75-1.00 is “excellent”). We have increased the threshold from 0.6 to 0.75 and updated our text as follows (including adding citations):

Any sites with intraclass correlations of less than 0.756 (a typical threshold for distinguishing between “fair” and “good”⁴² or “excellent”^{43,44} reliability) will be independently re-checked by Savage for another randomly selected song. If this is also less than 0.756, then all songs from that site will be checked and re-annotated until they achieve coefficients of at least 0.756.

Regarding your responses in the Author response document...

- Table 1 only had 4 minor edits, so it was not a “substantial expansion”. Did you intend to make substantial changes to this table?

We were referring to adding details about acoustic units to this table (since this was a key point brought up by two reviewers). We did not intend to make additional changes to this table.

- The video tutorial is great! I think a written protocol would also be handy for people who don’t want to scroll through the video if they have specific questions they want a refresher on.

by Corina Logan, 22 Apr 2025 06:46

Manuscript: https://doi.org/10.31234/osf.io/c2dba_v4

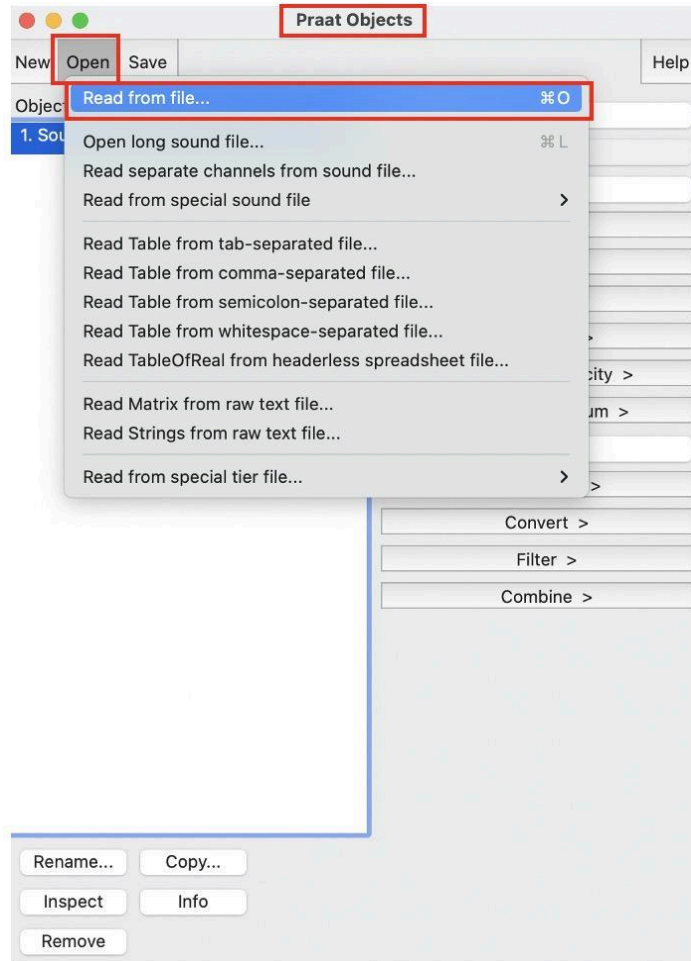
version: 4

We're glad you found it useful! We have supplemented with written details as suggested as follows:

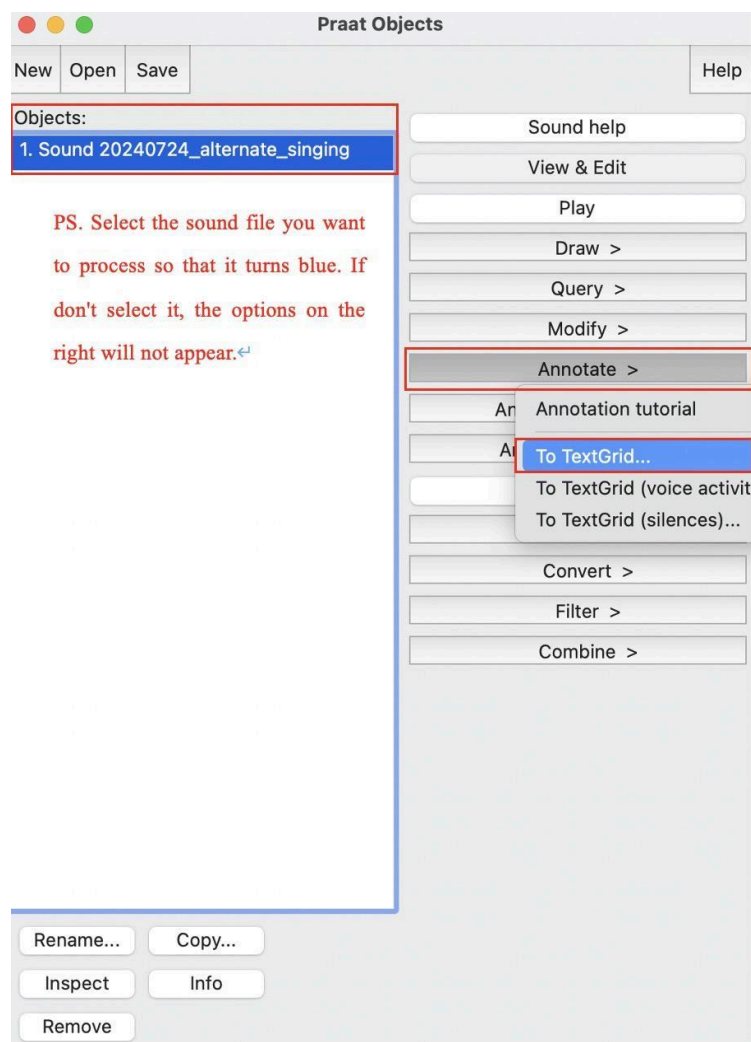
-A text version of the Praat tutorial follows:

A Tutorial for Doing Annotation in Praat

- 1. Download and install Praat if you haven't already.**
<https://www.fon.hum.uva.nl/praat/>
- 2. Prepare your audio recordings in a folder. Make sure your audio files are in .wav format.**
- 3. Open the audio recording you want to process:**
In Praat, go to *Praat Object > Open > Read from file...* , then select the file you want to work with.

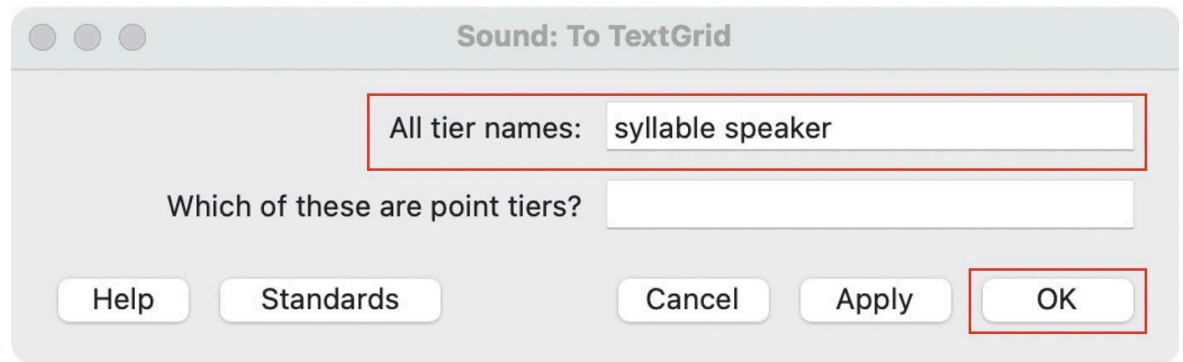


4. **Create a TextGrid for the corresponding audio file to make annotations:**
In the Objects window, select the audio file, then go to *Annotate > To TextGrid...*

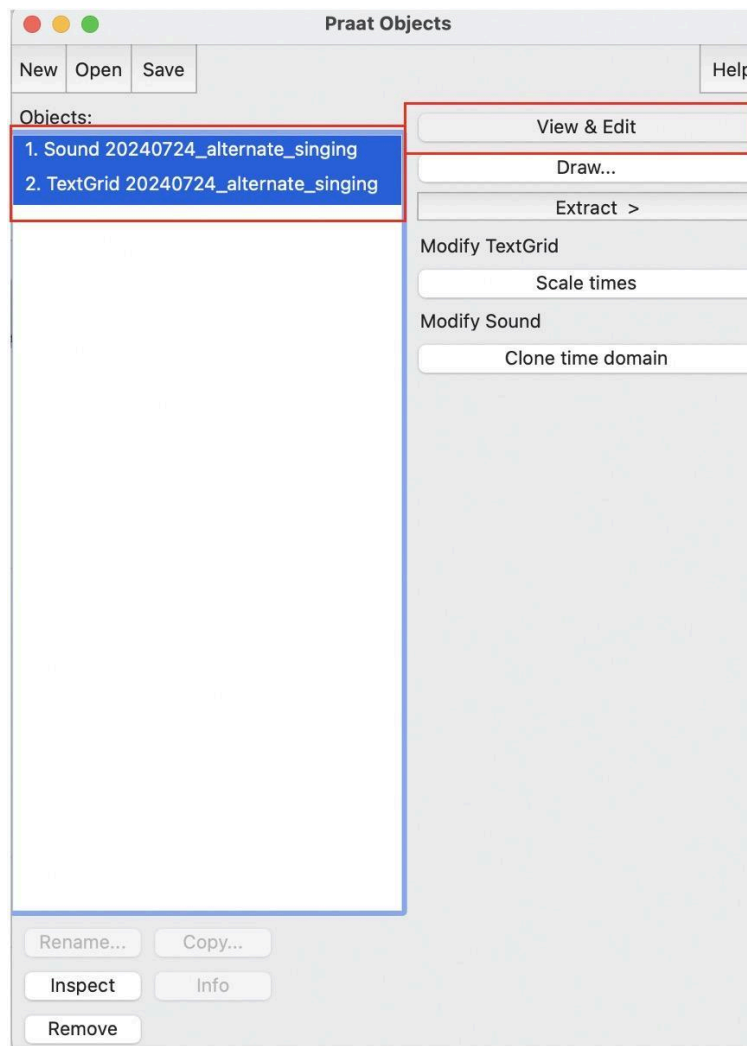


5. Then a small window called Sound: To TextGrid will appear. Clear all the default settings. In the All tier names: field, enter syllable speaker (if it is a song file, then enter note speaker), then click OK.

(A detailed explanation: We need to annotate two tiers: the first tier for syllables/notes, and the second tier for speakers. The logic for All tiers names is as follows: from left to right, list the first tier name, the second tier name, and then the third tier name.... separated by spaces. You are free to choose different tier names if you prefer; the tier names themselves do not affect data processing. They mainly serve as a guide to remind annotators what type of information should be labeled in each tier. Since we do not need to annotate any points, leave the point tiers empty.)

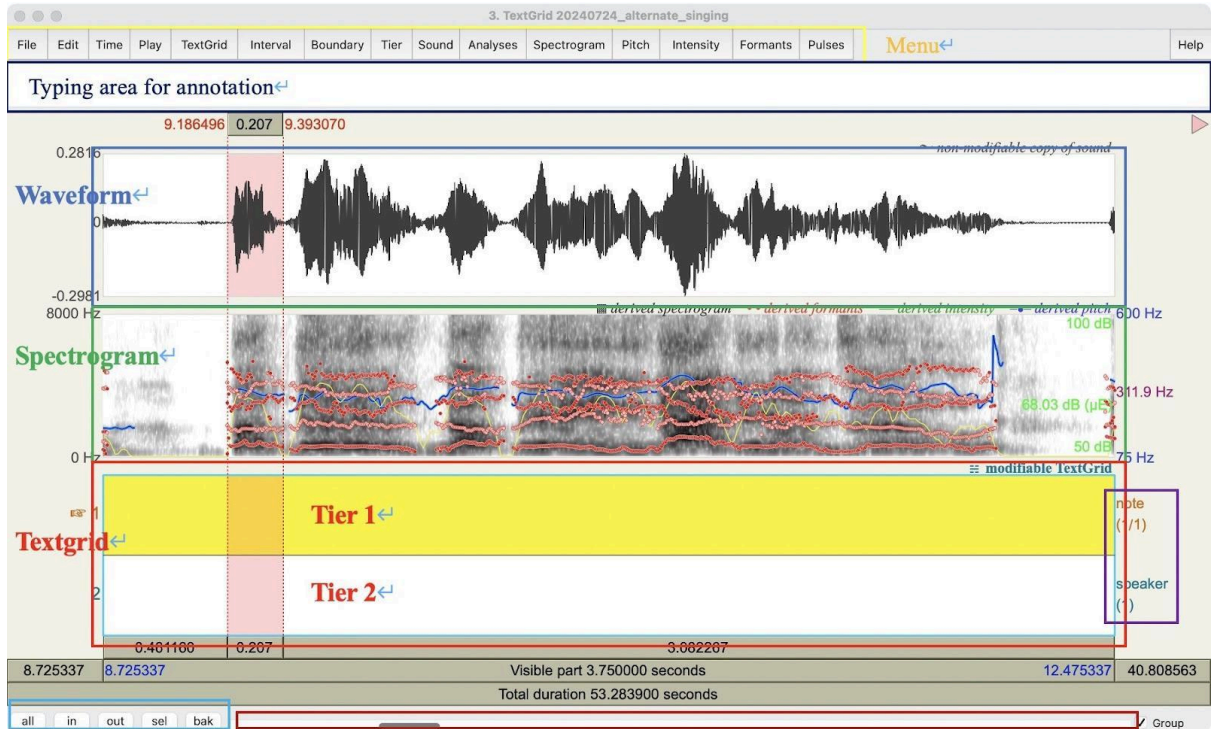


6. Now you have got a Sound file and a corresponding TextGrid file. Select both files, and click **View & Edit**.



7. Then the visualisation of combined sound and textgrid files appear. From top to bottom, you will see: the menu bar, typing area, waveform, spectrogram, TextGrid (tier 1 + tier 2), and time information.

(In the blue box at the bottom left corner below, "all" means to display the entire audio file within the window, "in" means to zoom in, "out" means to zoom out, and "sel" means to display only the selected regions. You can use these buttons together with the scrollbar on the right to adjust the view and select the appropriate region for annotation.)



8. **Add a boundary:** If you want to add a boundary, click on the waveform or spectrogram area at the position where you want the boundary to be. *(Do not click on the TextGrid area; clicking there is for adding or editing text labels.)*

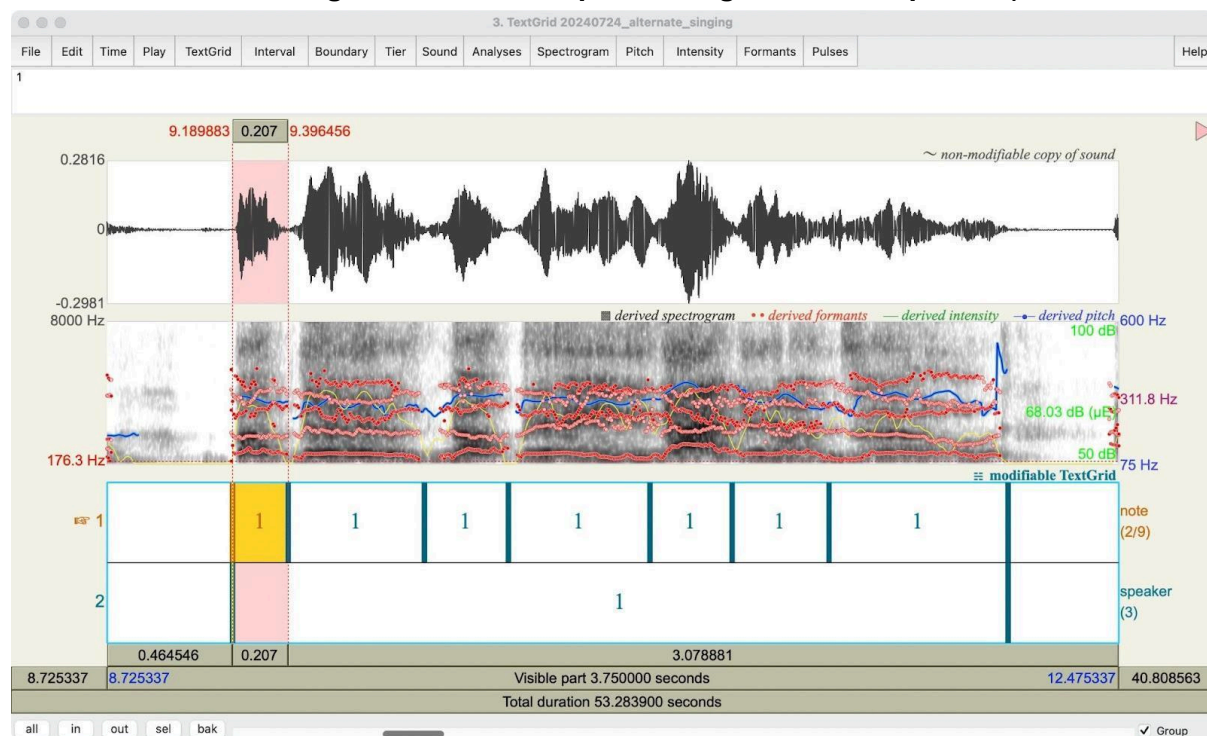
After clicking, a thick grey line will appear. Click the small circle  on top of the grey line *(clicking on the grey line itself won't work)*.

When the grey line turns blue, it means the boundary has been successfully added.

9. **Add annotation text:**
In the TextGrid area, select the region between two boundaries, and then type the text you want to add in the typing area.

(Since we ultimately need to know the duration of individual syllables or notes for different speakers, and also extract the complete speech

segments for each speaker, it's important to annotate speaker information. We recommend labeling based on Participant ID, e.g., 1, 2, 3,..., up to 10.)



10. Save the TextGrid file:

After completing all annotations, go back to the Praat Objects window, select the TextGrid file, and choose **Save > Save as text file....**

