

Annotation Instructions

Procedure Summary

First Annotator	Second Annotator
<p>Check the “audio to do” file on Box to see what files are assigned to you You will be assigned two files: one as first annotator, one as second annotator.</p>	
<p align="center">Complete Part 1: First Pass Annotation</p> <p>Upload your list of target time-windows to the “Ready for First Annotator” folder on Box File-naming convention: {wav file name}_targets_{annotator code} e.g. sad_boyz_2_multiTrack.wav matches sad_boyz_2_multiTrack_targets_ANam</p>	
<p>Wait until both you and the second annotator for the file have completed Part 1: First Pass Annotation, then inclusively combine the lists of target time-windows for the audio file.</p>	
<p>Complete Part 3: Mark Utterance Boundaries Save your unannotated TextGrid file and upload to the “Ready for Second Annotator” folder on Box</p> <p>File-naming convention: just use the default filename praat gives you - the exact same as the audio file name</p> <p>e.g. sad_boyz_2_multiTrack.TextGrid matches sad_boyz_2_multiTrack.wav</p>	
<p>Complete Part 4: Annotation Use your pre-segmented TextGrid</p>	<p>Complete Part 4: Annotation Use the first annotator's pre-segmented TextGrid</p>
<p align="center">When you are done with annotation:</p> <p>Upload your TextGrid to the individual folder for your audio file within the “Completed Annotations” folder on Box. File-naming convention: Keep the full filename of the audio file, but add the prefix: {annotator code}_ e.g. ANam_sad_boyz_2_multiTrack.TextGrid matches sad_boyz_2_multiTrack.wav</p>	
<p align="center">Email me to tell me you’ve finished a file!</p>	

Introduction and Project Overview

Welcome to the annotation team! The purpose of this annotation project is to create a corpus of utterance-length naturalistic speech samples that have been annotated for the presence/absence of irony. This corpus will be used for my dissertation research regarding irony prosody and machine learning approaches to irony recognition in speech data.

This document will give you an understanding of verbal irony, a description of my heuristic for identifying it based on conversational context, and procedures for systematic annotation and segmentation of long-form audio files.

What is (Verbal) Irony?

Definition

For the purposes of this study, verbal irony is a very broad term. My working definition is:

“Utterances in which the literal meaning of the speaker’s words is in opposition to, or at least different from, the speaker’s intended meaning - usually with some kind of humorous intent.”

In this sense, irony is an umbrella term that captures quite a few forms of “joking around” including, most notably, sarcasm, but also phenomena such as hyperbole, understatement, certain kinds of rhetorical questions, satire, and jocularly.

Examples of Irony in Conversation

Irony is highly context-dependent. For the sake of this study, irony will be identified within its conversational context and judged based on the response to the potentially ironic utterance by listeners.

Classic Examples

1. Sarcasm, hyperbole

- A: “Please don’t tell Susan I couldn’t eat the cake. It was sweet of her, but this is terrible...”
- B: “Yeah, sorry about that. **As we all know, she’s an excellent chef.**”
- [laughter]

2. Understatement

- A: [Entering from a blizzard, dressed in multiple layers of thick outdoors wear] “Brrr!”
- B: “**Little chilly today, huh?**”
- [laughter]

3. Rhetorical Question

- A: "Dude that test was easy! I finished in like, 10 minutes and I've never felt so confident about anything in my life"
- B: [feigning concern] "**Oh no, do you think you passed?**"
- A: [laughing] "Oh shut up."

Irony Chaining

Sometimes in conversation, one ironic utterance will be answered by another, and another and another, as speakers build upon the joke to, sometimes, construct a full fictional scenario.

- A: "I had a really great date last night. He had such a wonderful sense of humor!"
- B: "**But how can you be sure it wasn't just a clown in disguise?**"
- C: "**You know they look just like us without the makeup.**"
- B: "**Or you know how well they fit in tight places. It could have been six clowns in one regular-guy suit, just feeding each other lines.**"
- A: "**Well then I'll thank you to leave me and my six new clown lovers alone.**"
- [laughter]

Labeling Irony in Conversational Context

This method of labeling irony based on conversational response was developed following the example of Brown et al. (1980) and is intended to be an objective method for annotating irony. The following list of conversational responses to irony were distilled from the results of Gibbs (2000).

Conversational Responses to Irony

An utterance will be labeled as ironic if it is followed in conversation by any (combination) of the following:

1. Laughter (including laughter by the speaker of the utterance in question)
2. Response to ironic meaning
3. Ironic response (as in the "Irony Chaining" example)
 - a. This forms an "irony chain" that will need to be identified backward
 - b. The irony chain stops when the conversational topic changes.
4. In-situ explanation by speaker that irony was intended
 - a. E.g. The joke wasn't understood and the speaker says "Ah, sorry, that was a joke"

A note about laughter

Laughter within a conversation is not always loud and obvious. It could just be a quick chuckle at the end of a "yeah" or an "oh" or a "right." It could just be a quick exhale. If you can hear it and it sounds like laughter, use your best judgement.

Possible Near Misses

One very common near miss comes in the form of the example above where the speakers are collaborating to create and build-upon a fictional scenario for the purposes of irony/jocularity. This can sometimes be difficult to identify as irony, and may initially strike you as a false positive (see below), as there is not a clear “ironic meaning” and “literal meaning” to subsequent utterances after the first. Take the example below:

- A. “What’s the point of those little cocktail umbrellas anyway?”
B. **“They keep the cockroaches from getting too much sun on their beach days.”**
C. **“Awww, and they use the napkins as beach towels!”**
B. [nodding sagely] **“And the maraschino cherries are their watermelons.”**
[laughter]
A. **“Okay, and that all makes perfect sense, but how do they get to the beach? Do they fly, or do they have tiny cars, or...?”**
B and C. **“Toothpick hiking sticks!”**
[laughter]

In cases like this it is clear that the literal meaning of the words is not meant to be taken seriously, but the “ironic meaning” is not as clear as it would be in one of the “Classic Examples” of irony (see above). In this example, the first utterance by Speaker B can be explained as an ironic form of mockery or gentle critique. As in, the purpose of the cocktail umbrellas (simply as decoration) should be obvious, so a clearly-wrong, fantastical explanation is provided instead. What follows, building upon that initial ironic utterance, remains in the spirit of the jocularity of the initial utterance as the group of speakers adds new elements to the fictional world created by speaker B’s ironic answer.

Possible False Negatives

Sometimes irony isn’t acknowledged within conversational context. In these cases, under the current scheme, that ironic utterance would not be labeled as ironic, nor used as ironic data. However, it also should not be used as non-ironic data. If you hear an utterance that you have an instinct is ironic, but is not followed by an appropriate conversational response, make a note on the “Annotator Intuition” tier, and do not mark the utterance as either ironic or non-ironic.

Possible False Positives

Sometimes a non-ironic utterance may be followed by one of the conversational responses detailed in the Irony Labeling Guide. In these cases, you should rely on the definition of irony in this document to verify if a given utterance should be labeled as ironic or not.

1. Literally describing humorous events - may result in laughter without irony
 - A: “Yeah, people have a really hard time getting my name right over the phone. I’ve gotten Ellen, Heather, Hanna, Allen...”

- B: [laughing] "Allen???"
- A: "No lie. Allen."

Cases like this are funny, but the speakers are literally discussing a true state of affairs that happens to be funny. If speaker A had continued listing increasingly strange and fantastical name variations, irony could occur in the form of hyperbole, but it is not ironic to simply describe the true state of affairs.

2. Tripping over words/accidentally saying something funny - may result in laughter without irony.
 - A: "So I went to the store to buy some belly jeans
 - B: [chuckling] "What on earth are belly jeans?"
 - A: "Ah heck. I meant jelly beans! Jelly beans!"
 - [laughter]

Again, cases like this are funny for the conversational participants, but the unintentional nature of the initially funny utterance and subsequent utterance focused on correcting it preclude them from being irony. If speaker A had responded to speaker B's follow-up question with an ironic explanation of what belly jeans might be however, an ironic exchange could have grown out of this utterance. In either case, the initial mistake would not be considered ironic.

3. Identifying the beginning of an irony chain - sincere utterances with ironic responses. Examining the Irony Chaining example, we can see that the irony chain in question began with a non-ironic statement. While this is not required, it is also not uncommon for a non-ironic utterance to have an ironic response, despite such an ironic response potentially being a signal that the preceding utterance may have been ironic.

In cases such as this, it is important to remember this study's definition of irony. Is there a secondary meaning to the utterance that is **different** from the literal semantic content of the utterance?

Part 1: First-Pass Annotation

Although the data for this study was specifically selected to maximize the concentration of conversational irony, there will still likely be many more non-ironic utterances than ironic within an hour-long podcast episode. In order to target close annotation and reduce the gap between the amount of ironic and non-ironic data gathered while working as efficiently as possible, a first-pass step is necessary.

For this first pass, you may listen to the full episode at 1.5x-2.0x speed (episodes have been uploaded privately to Illinois Media Space to allow for this functionality), and roughly note the time-stamps of laughter. As discussed in the previous section, laughter is a relatively strong cue for irony. If you have an intuition that another time-point may be promising, you may note it as well. This step is simply to narrow down the amount of data considered for close annotation, so precision is not an absolute necessity.

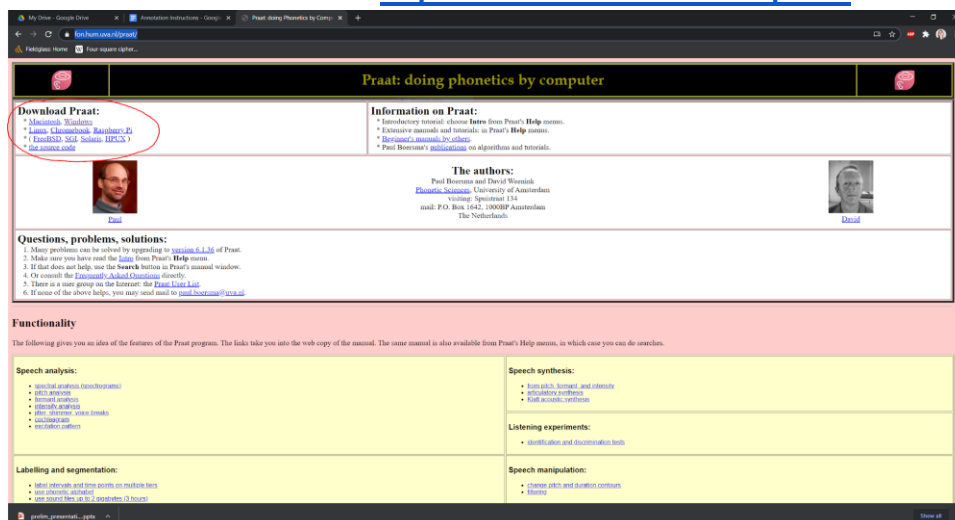
Once you have listened to the episode and noted promising time-stamps, you may find it helpful to convert these time-stamps to 30-second windows as a final step. You should consider (at minimum) the 15 seconds before the time-stamp you noted and the 15 seconds after for your close annotation. These time-windows indicate where you should focus your close annotation in Step 3, but if you find promising data beyond the pre-identified windows, you are welcome to annotate it as well.

An example first-pass output can be found in the file “First Pass Example (Sad Boyz 2)”.

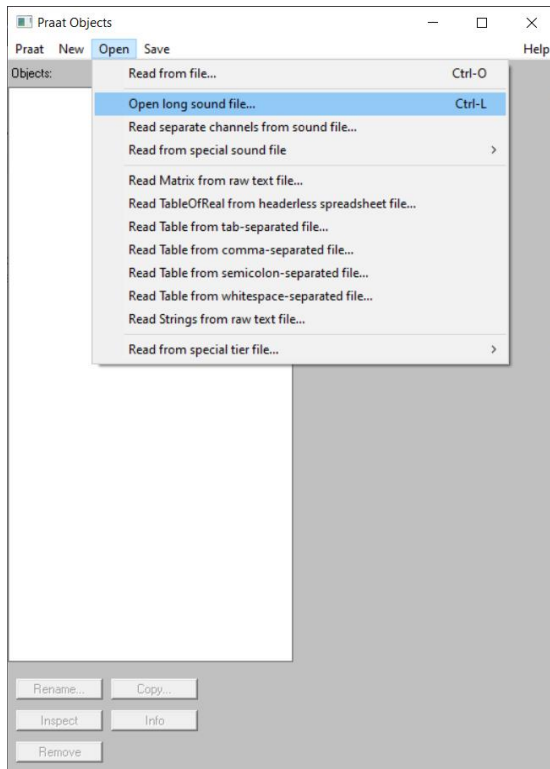
Note: Precision is not as big a priority at this point. You are just giving yourself some target areas to examine more closely and annotate.

Part 2: Getting Started with Praat

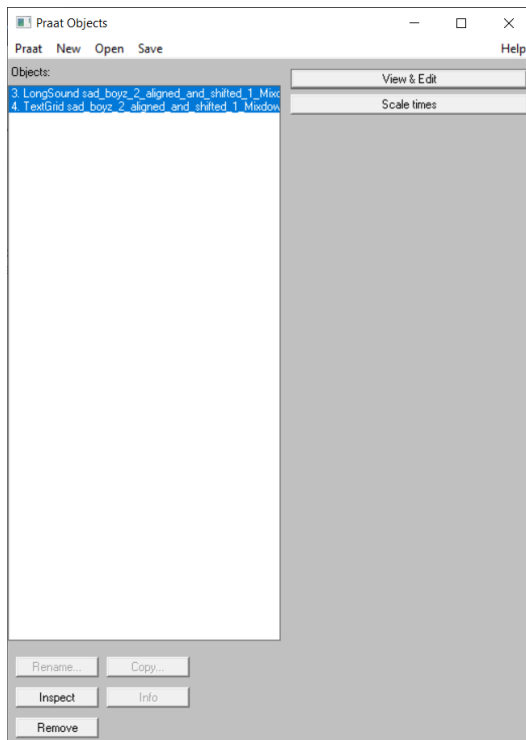
1. Download Praat for free at <https://www.fon.hum.uva.nl/praat/>



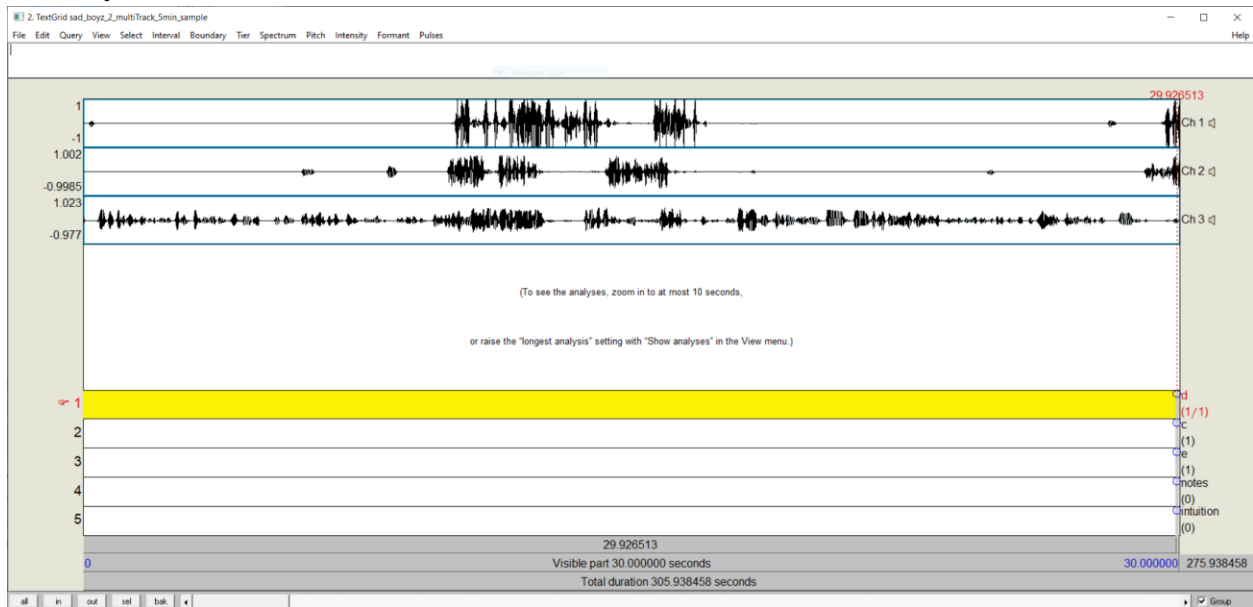
2. In the Praat Objects window, select “Open” then “Open long sound file...” or use Ctrl-L



3. Open the corresponding TextGrid file by selecting “Open” and “Read from file...”
4. In the Praat Objects window, select both the audio file and the TextGrid file, then click “View & Edit”



5. Zoom in to a manageable amount of visible sound. This is entirely up to your preference, but 7-15 seconds works well for me.



Part 3: Mark Utterance Boundaries (Annotator 1)

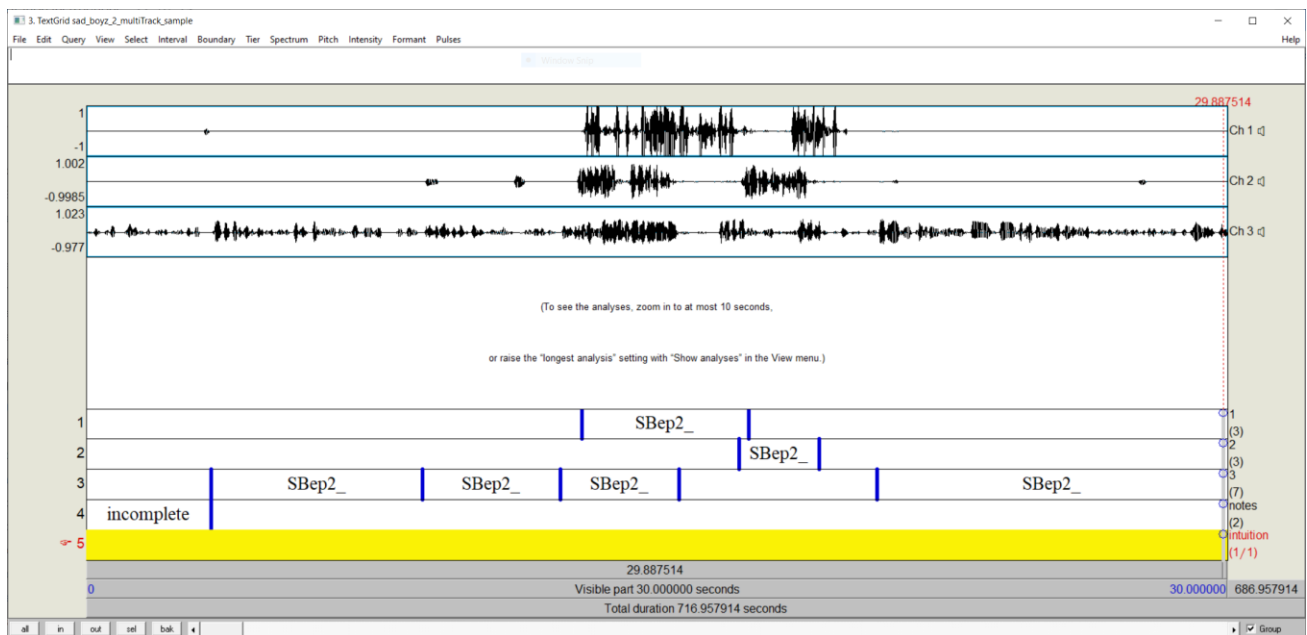
Procedure

- If you are Annotator 1 for a given long sound file, you should complete this step before moving to annotation.
- Mark utterance boundaries for **only** the relevant speaker's tier.
- For clarity as to which segments are utterances meant to be captured, add the beginning of the filename interval you have marked.
 - Add only the podcast name abbreviation and episode number (e.g. "SBep2_"). It will be the same for every utterance you mark.
- Once you have marked utterance boundaries for all the target sections of the file, use "Save TextGrid as text file" on the File menu to save your work.
- Finally, upload the blank, pre-segmented TextGrid to the "Ready for Second Annotator" folder on Box

Identifying Utterance Boundaries

- Utterances are usually conversational turns, so you'll place a boundary when a new speaker starts talking
- Sometimes a turn will contain more than one utterance. In this case, segment on major syntactic boundaries such as the end of a sentence.

- Utterance boundaries should also be marked when a speaker starts a “quote.” e.g. “and I was like, ‘I’ve never sent an email in my life!’” ← would have a break before “I’ve”
- Use your best judgement in cases where a speaker starts and then restarts an utterance, begins with an “um,” or uses filler words such as “um” and “and” to connect what could be multiple utterances into one very long utterance.
 - Whether these cases result in one long utterance connected with an “um,” or two shorter utterances with the “um” excluded, either outcome is fine.
- Use your best judgement when deciding whether to include utterances that are just “mmm” or “yeah” or “right” (e.g. “active listening” words).
- Try to trim off laughter at the end of utterances as much as possible.



Part 4: Annotation

There will be a tier for each speaker. Some episodes have two speakers, some have three. Your annotation in the tier for a given speaker will follow this format:

- Indicate the filename using the following format:
 - {1}ep{2}_{3}{4}-{5}
 - 1) Podcast name abbreviation (e.g. “Sad Boyz” → “SB”)
 - a) This will already be present from the procedure in Part 3
 - 2) Episode number (e.g. “2”)
 - a) This will already be present from the procedure in Part 3
 - 3) Speaker ID (e.g. “c”) – added programmatically
 - a) This will match the tier label

- 4) Utterance number - the utterance number **within the episode** (e.g. “1”) – added programmatically
- 5) Irony label (“N” for “non-ironic”, “I” for “ironic”)
- Full example: SBep2_c1-N

It is important that these filenames be uniform and precise. Please be accurate and consistent with case and other formatting!

Update on Speaker ID and utterance number

- After completing utterance segmentation (you still must manually add the episode prefix (e.g. “SBep2_”) to each utterance you identify), run add_stamps.py in the same directory your segmented TextGrid is in. This will automatically add speaker IDs and utterance numbers to each identified utterance.
- Please double check that the utterance numbers are correct and re-run add_stamps.py or ask for help if not. The utterance number indicates the utterance’s position within the entire episode and is not speaker dependent. They should begin at 1 and continue sequentially to n where n is the total number of utterances identified in the episode, irrespective of speaker.
- After completing this step, you may continue with annotation.

The “annotator intuition” tier

On this tier you may notate if you disagree with the outcome of applying the irony labeling scheme. Only use this tier if you disagree, and only indicate with “I” or “N” as in the “irony” tier.

Make sure that the irony label in the “filename” tier is based on your application of the irony labeling scheme as laid out earlier in this document

The “notes” tier

- Include any notes you have about why an utterance was marked as ironic or non-ironic.
- This is also where you may elaborate about disputes notated in the “annotator intuition” tier.

Segments to exclude

- If you, as first annotator, feel a segment should be excluded, do not mark its utterance boundaries or add the filename prefix.
 - If you are unsure, go ahead and mark its boundaries for the Part 3 step, then exclude it with notes in Part 4
- If you, as second annotator, feel a segment should be excluded, but the first annotator has marked it as an utterance, delete the filename prefix from the segment for the proposed utterance and leave it blank.

- ## Sample Annotated Selection



- Remove a boundary - alt+backspace
- Adding boundaries (works with segments!)
 - First tier - ctrl+f1
 - Second tier - ctrl+f2
 - Third tier - ctrl+f3
 - etc.

Saving your work - **Extremely Important!**

In order to save your work, you must use “Save TextGrid as text file” on the File menu. If you skip this step and close out of Praat, all your work will be lost!

When you return to the file, simply open the sound file again in the “Praat Objects” window, and open the text file you’ve already been working on. Then select both and click “View & Edit” as in Step 6 of Part 1.

After you’ve finished annotating a file

Once you have finished annotating all the target areas of an audio file, check the “Completed Annotations” folder on Box to see if a folder for that audio file already exists. If it doesn’t, create one.

In the individual folder for your audio file within the “Completed Annotations” folder, upload the full, long audio file itself (if it isn’t already there), and your TextGrid.

TextGrid naming convention: Keep the full filename of the audio file, but add a prefix of the format: **{annotator code}**_(e.g. ANam_sad_boyz_2_multiTrack.TextGrid)

Finally, email me to tell me you’ve finished a file!