

A problem named *Utterance*

Utterances can be phrases (one word or more) or whole sentences. For our Talk Together Study, we are interested in Mean Length of Utterance (MLU), and, therefore, will only consider utterances as being sentences (NOTE: a single word can be a sentence! E.g. “Eat!” - said as a command).

We look/listen for boundary markers: intonation and pauses. Listen out for falling intonation; there would be a pause immediately after (look at the waveform) and intuitively you would want to put a full-stop there. Listen out for rising intonation; there would be a pause immediately after (look at the waveform) and intuitively you would want to put a question mark there.

|| I bought um *satay* and *nasi lemak* ah and er ate them all up || - 1 utterance here

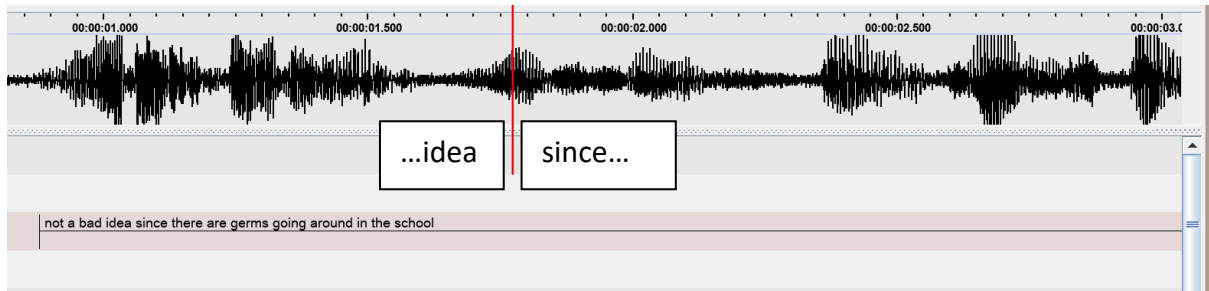
1. Look for complete sentences.
2. IF NO COMPLETE SENTENCES – we look for chunks that go together in intonation patterns. E.g. || I want to || || Hey did you eat the last piece? || - 2 utterances
3. An utterance can be as long as a complex sentence (e.g. “He went home, but did not have his keys.”)

Listen out for falling intonation—marks the end of a sentence—or a rising intonation for questions (this rise is more prominent than the one for a phrase). E.g.

	Utterance	Utterance	Utterance
Grandfather1 (Utterance) [117]	I said :d:ah OK :d:lah I just wash it :d:lah damn fast one :d:#!#?		
Grandfather1 (Language) [136]	English		
Grandfather1 (Chunk) [81]	I said :d:ah OK :d:lah I just wash it :d:lah damn fast one :d:#!#		
Grandfather1 (Translation) [13]			

(Note: I have not heard the audio for the above. The space between the utterances is where I would have a longer pause if I were the speaker. In addition, 'I said <interjection: ah>' is not a complete sentence and lacks semantic sense in the above.)

The whole thing below is a single utterance.



Different corpus teams (including non-linguistic ones) have pause rules ranging from 0.5 seconds to 2 seconds of duration. These are arbitrary rules. Pauses are relative to speakers' rate of speech (Demol, Verhelst, & Verhoeve, 2007).

Proposed solution: pause + intonation + syntactical structure (i.e. that which constitutes a sentence by native speakers)

Interrupted speech

|| I am such a *glutton* ||

|| Hey, I am no *pig* ||

|| Yes, what a greedy *pi...* || - still considered an utterance

Reference

Demol, M., Verhelst, W., & Verhoeve, P. (2007). The duration of speech pauses in a multilingual environment. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 1(1)*, 117–120.