Speech Rate and Articulation Rate

Speech rate is usually measured in syllables per minute or syllables per second. LaBB-CAT can calculate this if it has the following information:

1. start and end times of utterances, from which the each utterance duration can be calculated (usually, transcripts you upload to LaBB-CAT include this information), and
2. the number of syllables in each word token, which can be obtained either

   - by using the CELEX layer manager to tag each with with its syllable count, or
   - if the data has been force-aligned, using a lexicon with syllabification information, the syllables themselves can be reconstructed.

If the speech has not been force-aligned, LaBB-CAT only knows, for each utterance, how many syllables were uttered between the start and end times of the utterance; any inter-word pauses during the utterance are counted as speech. Usually, this level of granularity is referred to as the *Speech Rate*.

If the speech has been force-aligned, LaBB-CAT can use the start and end times of the individual word tokens, and so can exclude inter-word pauses from its rate calculation. This higher level of granularity is usually referred to as the *Articulation Rate*.

The rate annotations can cover a number of different scopes with different granularities, e.g.

- the local rate for each utterance,
- the local rate for each speaker turn (which may include many utterances),
- the global rate for the entire recording, or
- the global rate for the speaker, across all the recordings they appear in.

Once the rate annotations are generated, they can be searched or extracted, in the same way other annotations can.

Computing Speech/Articulation Rate

1. Create a *word layer* that tags each token with the number of syllables in the word. For example, if you've got English data, and have the CELEX Layer Manager installed, you can achieve this by creating a new word layer with the following characteristics:

   - *Layer ID*: `syllableCount`
   - *Type*: *Number*
   - *Alignment*: *None*

- *Manager*: *CELEX English*
  …configured with the *Syllable count* option ticked
  IMPORTANT: Ensure tick the *First match only* option for this layer, as some words have more than one possible syllable count, and we don't want any words counted more than once.

Source layer: orthography ∨

Language: en.*

Delimiters:

Lookup Definition:

Presets

CELEX SQL Query

```
SELECT CHAR_LENGTH(PhonStrsDI
 FROM cxen_wordformphonologyp
 INNER JOIN cxen_wordformorth
 ON cxen_wordformphonologypro
 INNER JOIN cxen_wordform
 ON cxen_wordformphonologypro
 WHERE cxen_wordformortho.Wor
 ORDER BY cxen_wordform.cob D
```

- ◯ Phonology
- ⦿ Syllable Count
- ◯ Morphology
- ◯ Syntax
- ◯ Lemma
- ◯ Frequency (wordform)
- ◯ Frequency (lemma)
- ◯ Syllables from Phonology
- ☑ First match only

Test word: testing

[Save] 2

.
If you do not use the CELEX Layer Manager, you may find the layer manager you use for phonemic transcriptions also has a similar *syllable count* option.

2. Once the layer has finished generating, have a look at a transcript or two to check the results. Each word should be tagged with a number corresponding to the number of

syllables:

```
1   2  1 1   1   1   1       3     1   1    1   1 1   1   1 1 1 1   2     1
um after a big night in town probably the one night of the year that I did go into town

1   1    2    3     2    1   1    1     1     1   1  1 1 1 1 1   1    1    1
um and very quietly opened the door and closed the door so as not to ah wake my wife .

1   1   1  1   1     1   1      2     1   1   1   1     3     1   2   1  1
um . and then of course once the earthquake hit . um - she most certainly was awake um and
```

3. Click the *phrase layers* menu option.
4. Add a new phrase layer for speech rate. Key points are:

   • The layer manager to use is the *Statistics Layer Manager*
   • The layer to summarize should be the *syllableCount* layer
   • The statistic to compute is *Label-Sum Rate (per minute)*, because we want LaBB-CAT to take the sum of all the labels, and then compute the rate from the start/end time.

Layer to Summarize: syllable count ˅

Statistic Label Sum Rate (per minute) ˅

Pattern to Match Regular expression

Context [⌃⌄] seconds

Pause threshold [⌃⌄] ms

☐ Main participant utterances only

Scopes [˅]

☑ Turns

☐ Utterances

☐ Participant in Transcript

Transcripts: [none] ˅

Participant: [add new attribute called speech rate] ˅

☐ Corpora

Transcript Types ☑ interview

Save

You can select difference scopes for the computation. The illustration above computes a local rate for each speaker *turn*, and a global rate for each *participant*.

5. Have a look in a transcript or two, and a participant or two, to see what the annotations you just generated look like.

Aligned Words

If forced alignment has already be done, the steps to compute speech rate are:

1. Create a *word layer* that reconstructs the syllables, using these steps.
2. Once the layer has finished generating, have a look at a transcript or two to check the results. Each word should be tagged with a phonemic transcription, separated into syllables:

'aɪ 'wɒz 'skræ blɪŋ 'ʌp ð 'wɔːl 'traɪ ɪŋ t 'gɛt 'aʊt   'θru: 'ði: 'wɔːd rəʊb ɪn 'stɛd 'ɒv   ð   'dɔː:
 I was scrabbling up the wall trying to get . out through the wardrobe instead of the . door

3. Select the *phrase layers* menu option.
4. Add a new phrase layer for speech rate. Key points are:

   • The layer manager to use is the *Statistics Layer Manager*
   • The layer to summarize should be the *syllable* layer
   • The statistic to compute is *Token Rate (per minute)*, because we want LaBB-CAT to take the count the number of syllables, and then compute the rate from the start/end time of each word.

Layer to Summarize:	syllable
Statistic	Token Rate (per minute)
Pattern to Match	Regular expression
Context	seconds
Pause threshold	ms

☐ Main participant utterances only

Scopes	
	☐ Turns
	☑ Utterances
	☐ Participant in Transcript
	Transcripts: [none]
	Participant: [add new attribute called articulation rate]
	☐ Corpora
Transcript Types	☑ interview

Save

You can select difference scopes for the computation. The illustration above computes a local rate for each *utterance*, and a global rate for each *participant*.

5. Have a look in a transcript or two, and a participant or two, to see what the annotations you just generated look like.

328.9817

I was scrabbling up the wall trying to get . out through the wardrobe instead of the . door