

Syllable rate vs. segment rate in perceived speech rate

Yahya Aldholmi

Department of Linguistics, King Saud University, Saudi Arabia
<https://doi.org/10.36505/ExLing-2020/11/0005/000420>

Abstract

This study examines the impact of differences in segment rate, which do not correspond to differences in syllable rate, on perceived speech rate in an unfamiliar language. Japanese speakers who were unfamiliar with Arabic judged the speech rate of Arabic stimuli with two levels of syllable complexity on a 1-7 scale. The participants perceived the complex syllable stimuli as being faster than the simple syllable stimuli. This is interpreted as evidence that differences in segment rate *do* influence speech rate perception. Hence, the significance of syllable rate, rather than of segment rate, for perceiving speech rate in some previous studies was likely due to extraneous factors associated with the stimuli and participants in those studies.

Keywords: syllable rate, segment rate, syllable complexity, speech rate, perception

Introduction

Previous studies have implemented various metrics to quantify speech rate, with syllable rate being the most commonly used (e.g., Aldholmi, Park, 2019), followed by segment rate (e.g., Seifart et al., 2018), among other measures. Two utterances with nearly the same syllable rate may have two different segment rates, and vice versa, due to differences in the respective languages' syllable structures. For instance, in the two English and Japanese six-segment words /stɹɪkt/ "strict" and /kisoku/ "rules", respectively, English incorporates all six segments into one syllable and hence exhibits a high segment rate *per syllable* (i.e., syllable complexity), while Japanese distributes the six segments over three syllables and hence exhibits a lower segment rate. Assuming speech duration is roughly the same (500s sec) for the two words, the speech rate would be 12 seg/sec (or 83.33 sec/seg) for both, but 2 syl/sec (or 500 sec/syl) for the English word and 6 syl/sec (or 166.66 sec/syl) for the Japanese word.

Considering such discrepancies, a few recent studies have attempted to *directly* investigate the impact of syllable complexity on perceived speech rate with varying results. For instance, syllable rate substantially affects speech rate perception for English natives (e.g., Plug & Smith, 2018), while segment rate (i.e., mora count) impacts speech rate perception for Finnish natives (e.g., O'Dell & Nieminen, 2019). More recently, segment rate was found to be related to stimulus duration in English. That is, "among stretches of speech

with various durations... stretches with higher segment rates... are judged as faster than stretches with lower segment rates” (Plug et al, 2020, p. 9).

Language familiarity may have played a role in these findings, because both listener-related factors such as phoneme restoration in language listening and language-specific characteristics such as syllable structure and mora rhythm could influence listeners’ estimates of speech rate. Therefore, the current study seeks to suppress these potential effects by recruiting speakers who have minimal/no experience with the language used in the experimental stimuli to investigate syllable complexity in speech rate perception. If segment rate (again, syllable complexity) is relevant to speech rate perception, listeners should be sensitive to the manipulated segment rate of unfamiliar speech presented to them. They should judge stimuli with complex syllables as being faster than stimuli with simple syllables. The present findings support this hypothesis.

Methodology

Stimuli

The stimuli consisted of complete Arabic sentences. Arabic root-pattern morphology allowed us to manipulate and construct stimuli with two levels of syllable complexity, simple and complex. Set 1 of the stimuli consisted of twenty-four six-word sentences with simple syllable structure ($Mean \approx 1$ C/syl) and a normal to slightly slow speech rate ($Mean \approx 3.9$ syl/sec). Set 2 of the stimuli was identical to Set 1 in terms of the number of sentences and nearly identical in terms of the number of words per sentence. However, the segment rate in Set 2 was modestly increased, resulting in a relatively higher syllable complexity ($Mean \approx 1.75$ C/syl) and hence the speech rate was relatively slower ($Mean \approx 3$ syl/sec). An Arabic native recorded the sentences using the first recorded sentence as a model for the target speech rate of the remaining sentences. Then, before running the experiment, five Arabic natives were asked to judge the speech rate of the stimuli on a 1-7 scale (1 = extremely slow, 4 = neutral, 7 = extremely fast) in order to make certain that stimuli speech rate was just as intended (not extremely fast/slow). They generally found all sentences to have a neutral (4) or slightly slow (3) speech rate ($Mean \approx 3.80$).

Procedure

Twenty-one male and female Japanese natives participated in a perception experiment. All participants had either no previous exposure to Arabic or, at most, knew a few simple words such /salaam/ “greetings”. The age of the participants ranged from 21 to 39, and none of them reported any hearing loss or impairment. They were introduced to the experiment and provided with five trials for practice. The experimental stimuli were presented in an auditory form and in a randomized order. The participants were asked to make estimates about the speech rate, using the same scale used by the original Arabic judges.

Results and discussion

As illustrated in Figure 1, the participants generally judged the complex syllable stimuli as being faster (*Mode* = 3, *Mean* = 3.23) than the simple syllable stimuli (*Mode* = 2, *Mean* = 2.21). The participants judged almost 50% of the simple syllable stimuli as being slow and almost 50% of the complex syllable stimuli as being slightly fast. They also judged approximately 20% of the simple syllable stimuli and only about 2% of the complex syllable stimuli as being extremely slow, while they judged slightly more than 18% of the complex syllable stimuli and slightly less than 2% of the simple syllable stimuli as being neutral.

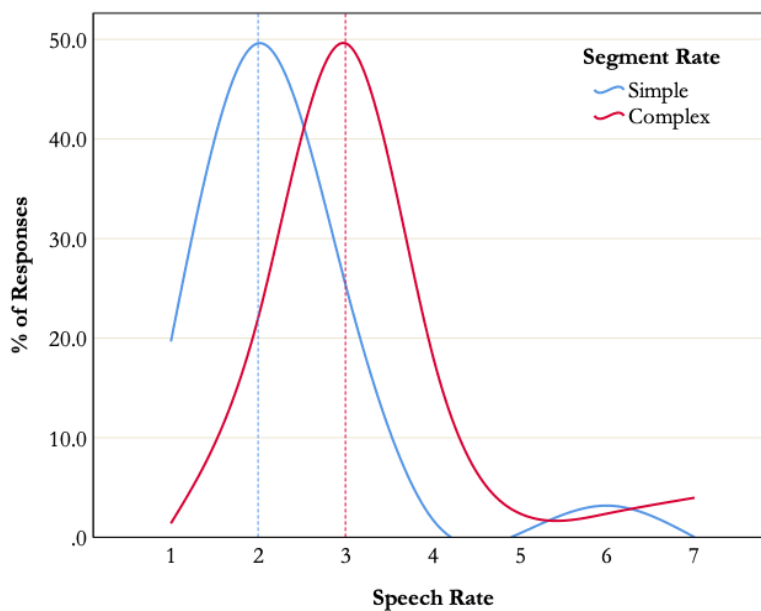


Figure 1. Proportional distribution of ratings of simple vs. complex stimuli

A repeated-measure regression model was performed with speech rate response as a predicted variable and segment rate/syllable complexity as a predictor variable. The model revealed that syllable complexity was an explanatory factor, Wald $\chi^2(1) = 147.99$, $p = 0.001$.

The results provide evidence that that segment rate does influence listeners' perception of speech rate. The difference in segment rate between the simple rate stimuli and the complex syllable stimuli is not that large (1 vs. 1.75 C/syl, respectively), but this triggered a one-point increment in the ratings, both with respect to the means and modes of the data. Syllable rate (# of syl) cannot be the reason for the high ratings in the complex syllable stimuli because it was lower than that in the simple syllable stimuli. Nor can stimuli durations be a

cause, as they were roughly equal across stimuli at both levels of syllable complexity; in fact, the simple syllable stimuli were slightly shorter in duration.

Thus, the findings are in agreement with O'Dell and Nieminen (2019) who found that listeners are sensitive to segment rate, but are in (at least partial) disagreement with other findings which claim that segment rate does not play a crucial role in speech rate perception. Specifically, the pattern in the current findings is at odds with Plug and Smith's conclusion that there is "no evidence that listeners are using segment rate to influence their tempo judgements" (2018, p. 282). In contrast to Plug et al (2020), the influence of segment rate was not contingent on durational differences in the present study, as duration was controlled across stimuli to be as equivalent as possible. Indeed, if a slight difference in stimuli length were to have had any effect in the present study, it should have triggered higher ratings in the simple syllable stimuli (as in Plug et al, 2020) because they were marginally longer. However, this was not the case.

Conclusion

The findings show that differences in segment rate influence the perceived speech rate even if the difference in segment rate is not accompanied by a difference in syllable rate. This diverges from previous studies, possibly because the listeners were presented with stimuli from an unfamiliar language. Information load, syntactic parsing, and semantic processing did not bias the listeners' estimates; the listeners were found to judge the speech rate based on syllable rate and segment rate, with complex segment rates triggering higher ratings. Future studies may increase the complexity of the experiment design by adding conditions to vary syllable rate and segment rate in unfamiliar speech.

References

- Aldholmi, Y., Park, H. 2019. Interactions between temporal acoustics and indexical information in speech rate perception. In *Perspectives on Arabic Linguistics XXXI: Papers from the annual symposium on Arabic Linguistics*, Norman, Oklahoma, 2017 (Vol. 8). John Benjamins Publishing Company.
- O'Dell, M., Nieminen, T. 2019. Syllable rate, syllable complexity and speech tempo perception in Finnish. In *Proceedings of the 19th International Congress of Phonetic Sciences 622-626*, Melbourne, Australia.
- Plug, L., Smith, R. 2018. Segments, syllables and speech tempo perception. In *Proceedings of the 9th International Conference on Speech Prosody*, 279-283, Poznan, Poland.
- Plug, L., Lennon, R., Smith, R. 2020. Listeners' sensitivity to syllable complexity in speech tempo perception. In *Proceedings of the 10th International Conference on Speech Prosody 2020*. Tokyo, Japan.
- Seifart, F., Strunk, J., Danielsen, S., Hartmann, et al. 2018. Nouns slow down speech across structurally and culturally diverse languages. *Proceedings of the National Academy of Sciences*, 115(22), 5720-5725.